

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 29, 2003, 10:46:52 ; Search time 5.67273 Seconds  
(without alignments)  
46.243 Million cell updates/sec

Title: US-09-707-738-22

Perfect score: 56

Sequence: 1 AKXVAWTLKAA 13

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	94.6	13	9 US-09-949-375A-18	Sequence 18, Appl
2	53	94.6	13	9 US-09-785-215-19	Sequence 19, Appl
3	53	94.6	13	10 US-09-894-018-69	Sequence 69, Appl
4	53	94.6	24	10 US-09-983-019-10	Sequence 10, Appl
5	53	94.6	27	10 US-09-983-019-4	Sequence 4, Appl
6	53	94.6	27	10 US-09-983-019-7	Sequence 7, Appl
7	53	94.6	55	9 US-09-949-375A-12	Sequence 12, Appl
8	53	94.6	80	10 US-09-894-018-97	Sequence 97, Appl
9	53	94.6	98	10 US-09-894-018-101	Sequence 101, Appl
10	53	94.6	106	10 US-09-894-018-95	Sequence 95, Appl
11	53	94.6	107	10 US-09-894-018-93	Sequence 93, Appl
12	53	94.6	107	10 US-09-894-018-103	Sequence 103, Appl
13	53	94.6	123	10 US-09-894-018-109	Sequence 109, Appl
14	53	94.6	130	10 US-09-894-018-99	Sequence 99, Appl
15	53	94.6	144	10 US-09-894-018-129	Sequence 129, Appl
16	53	94.6	147	10 US-09-894-018-131	Sequence 131, Appl
17	53	94.6	148	10 US-09-894-018-127	Sequence 127, Appl
18	53	94.6	157	10 US-09-894-018-117	Sequence 117, Appl
19	53	94.6	168	10 US-09-894-018-115	Sequence 115, Appl

20	53	94.6	169	10 US-09-894-018-119	Sequence 119, Appl
21	53	94.6	180	10 US-09-894-018-125	Sequence 125, Appl
22	53	94.6	206	10 US-09-894-018-111	Sequence 111, Appl
23	53	94.6	219	10 US-09-894-018-113	Sequence 113, Appl
24	53	94.6	255	10 US-09-894-018-89	Sequence 89, Appl
25	53	94.6	277	10 US-09-894-018-81	Sequence 81, Appl
26	53	94.6	280	10 US-09-894-018-83	Sequence 83, Appl
27	53	94.6	281	10 US-09-894-018-91	Sequence 91, Appl
28	53	94.6	308	10 US-09-894-018-105	Sequence 105, Appl
29	53	94.6	308	10 US-09-894-018-107	Sequence 107, Appl
30	53	94.6	340	10 US-09-894-018-137	Sequence 137, Appl
31	53	94.6	419	10 US-09-894-018-123	Sequence 123, Appl
32	53	94.6	585	10 US-09-894-018-87	Sequence 87, Appl
33	49	87.5	11	10 US-09-756-983-7	Sequence 7, Appl
34	49	87.5	11	10 US-09-828-574-12	Sequence 12, Appl
35	36	64.3	97	10 US-09-873-880-14	Sequence 14, Appl
36	36	64.3	296	10 US-09-873-880-34	Sequence 34, Appl
37	36	64.3	296	10 US-09-873-880-36	Sequence 36, Appl
38	35	62.5	479	9 US-09-738-626-3765	Sequence 3765, Ap
39	34	60.7	170	9 US-09-479-040-3	Sequence 3, Appl
40	34	60.7	409	9 US-09-738-626-4552	Sequence 4552, Ap
41	34	60.7	489	9 US-09-888-320-2	Sequence 2, Appl
42	34	60.7	579	9 US-09-922-364A-19	Sequence 19, Appl
43	34	60.7	579	9 US-09-254-590-19	Sequence 19, Appl
44	34	60.7	579	9 US-10-115-695-19	Sequence 19, Appl
45	34	60.7	580	9 US-09-922-364A-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-949-375A-18  
; Sequence 18, Application US/09949375A  
; Patent No. US20020172673A1  
; GENERAL INFORMATION:  
; APPLICANT: KLYSNER, Steen et al.  
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING ICE  
; FILE REFERENCE: 3631-0111P  
; CURRENT APPLICATION NUMBER: US/09/949,375A  
; CURRENT FILING DATE: 2002-01-18  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.  
US-09-949-375A-18

Query Match 94.6%; Score 53; DB 9; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0016;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12  
|||  
Db 1 AKFVAWTLKAA 12

RESULT 2  
US-09-785-215-19  
; Sequence 19, Application US/09785215  
; Publication No. US20020187157A1  
; GENERAL INFORMATION:  
; APPLICANT: JENSEN, Martin Roland et al.  
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID  
; FILE REFERENCE: 3631-0107P  
; CURRENT APPLICATION NUMBER: US/09/785,215  
; CURRENT FILING DATE: 2001-02-20  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19

```
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial T-cell epitope capable of binding to a large portion
; OTHER INFORMATION: of MHC Class II molecules in a variety of animals
US-09-785-215-19

Query Match          94.6%; Score 53; DB 9; Length 13;
Best Local Similarity 91.7%; Pred. No. 0.0016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
   |||||
Db 1 AKFVAAWTLKAA 12

RESULT 3
US-09-894-018-69
; Sequence 69, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chesnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69

Query Match          94.6%; Score 53; DB 10; Length 13;
Best Local Similarity 91.7%; Pred. No. 0.0016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
   |||||
Db 1 AKFVAAWTLKAA 12

RESULT 4
US-09-983-019-10
; Sequence 10, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (3)..(3)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-10

Query Match          94.6%; Score 53; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
   |||||
Db 1 AKXVAAWTLKAA 12

RESULT 5
US-09-983-019-4
; Sequence 4, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (6)..(6)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-4

Query Match          94.6%; Score 53; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
   |||||
Db 4 AKXVAAWTLKAA 15

RESULT 6
US-09-983-019-7
; Sequence 7, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
```

; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 7  
 ; LENGTH: 27  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1..7)  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (17) (17)  
 ; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine  
 US-09-983-019-7

Query Match 94.6%; Score 53; DB 10; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.0033;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 15 AKXVAAWTLKAA 26

RESULT 7  
 US-09-949-375A-12  
 ; Sequence 12, Application US/09949375A  
 ; Patent No. US20020172673A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KLYSNER, Steen et al.  
 ; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE  
 ; FILE REFERENCE: 3631-011P  
 ; CURRENT APPLICATION NUMBER: US/09/949,375A  
 ; CURRENT FILING DATE: 2002-01-18  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 12  
 ; LENGTH: 55  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.  
 US-09-949-375A-12

Query Match 94.6%; Score 53; DB 9; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.0065;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 13 AKFVAAWTLKAA 24

RESULT 8  
 US-09-894-018-97  
 ; Sequence 97, Application US/09894018  
 ; Patent No. US20020119127A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIMMUNE, Inc.  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Chestnut, Robert  
 ; APPLICANT: Livingston, Brian  
 ; APPLICANT: Baker, Denisw  
 ; APPLICANT: Newman, Mark  
 ; APPLICANT: Brown, David  
 ; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
 ; FILE REFERENCE: 39963-20033.00  
 ; CURRENT APPLICATION NUMBER: US/09/894, 018  
 ; CURRENT FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: PCT/US00/35568  
 ; PRIOR FILING DATE: 2000-12-28  
 ; PRIOR APPLICATION NUMBER: US 60/173,390

; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: US 60/284,221  
 ; PRIOR FILING DATE: 2001-04-16  
 ; NUMBER OF SEQ ID NOS: 368  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 97  
 ; LENGTH: 80  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV.3s2(-3)  
 US-09-894-018-97

Query Match 94.6%; Score 53; DB 10; Length 80;  
 Best Local Similarity 91.7%; Pred. No. 0.0094;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 32 AKFVAAWTLKAA 43

RESULT 9  
 US-09-894-018-101  
 ; Sequence 101, Application US/09894018  
 ; Patent No. US20020119127A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIMMUNE, Inc.  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Chestnut, Robert  
 ; APPLICANT: Livingston, Brian  
 ; APPLICANT: Baker, Denisw  
 ; APPLICANT: Newman, Mark  
 ; APPLICANT: Brown, David  
 ; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
 ; FILE REFERENCE: 39963-20033.00  
 ; CURRENT APPLICATION NUMBER: US/09/894,018  
 ; CURRENT FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: PCT/US00/35568  
 ; PRIOR FILING DATE: 2000-12-28  
 ; PRIOR APPLICATION NUMBER: US 60/173,390  
 ; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: US 60/284,221  
 ; NUMBER OF SEQ ID NOS: 368  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 101  
 ; LENGTH: 98  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV.PC3  
 US-09-894-018-101

Query Match 94.6%; Score 53; DB 10; Length 98;  
 Best Local Similarity 91.7%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 34 AKFVAAWTLKAA 45

RESULT 10  
 US-09-894-018-95  
 ; Sequence 95, Application US/09894018  
 ; Patent No. US20020119127A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIMMUNE, Inc.  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Chestnut, Robert  
 ; APPLICANT: Livingston, Brian

```
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s1
US-09-894-018-95
```

```
Query Match          94.6%; Score 53; DB 10; Length 106;
Best Local Similarity 91.7%; Pred. No. 0.012; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;
```

```
Oy 1 AKXVAAWTLKAA 12
||| |||||
Db 32 AKFVAAWTLKAA 43
```

```
RESULT 11
US-09-894-018-93
; Sequence 93, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s1
US-09-894-018-93
```

```
Query Match          94.6%; Score 53; DB 10; Length 107;
Best Local Similarity 91.7%; Pred. No. 0.012; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;
```

```
Oy 1 AKXVAAWTLKAA 12
||| |||||
Db 32 AKFVAAWTLKAA 43
```

```
RESULT 12
US-09-894-018-103
; Sequence 103, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.PC4
US-09-894-018-103
```

```
Query Match          94.6%; Score 53; DB 10; Length 107;
Best Local Similarity 91.7%; Pred. No. 0.012; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;
```

```
Oy 1 AKXVAAWTLKAA 12
||| |||||
Db 33 AKFVAAWTLKAA 44
```

```
RESULT 13
US-09-894-018-109
; Sequence 109, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Artificial Sequence
```



```
;
; FEATURE:
; OTHER INFORMATION: AOSI.K
US-09-894-018-109

Query Match      94.6%; Score 53; DB 10; Length 123;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKAA 12
Db 34 AKFVAAWTLKAA 45

RESULT 14
US-09-894-018-99
; Sequence 99, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s3
US-09-894-018-99

Query Match      94.6%; Score 53; DB 10; Length 130;
Best Local Similarity 91.7%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKAA 12
Db 32 AKFVAAWTLKAA 43

RESULT 15
US-09-894-018-129
; Sequence 129, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
```

```
;
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BCL A2 #88
US-09-894-018-129

Query Match      94.6%; Score 53; DB 10; Length 144;
Best Local Similarity 91.7%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKAA 12
Db 34 AKFVAAWTLKAA 45

Search completed: January 29, 2003, 10:59:00
Job time : 6.67273 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:37 ; Search time 9.45455 Seconds  
(without alignments)  
40.457 Million cell updates/sec

Title: US-09-707-738-22  
Perfect score: 56  
Sequence: 1 AKXVAATLKAAX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	94.6	13	1	US-08-305-871A-22
2	53	94.6	13	1	US-08-305-871A-23
3	53	94.6	13	4	US-08-788-822A-26
4	53	94.6	13	4	US-08-788-822A-27
5	53	94.6	25	3	US-09-075-257A-17
6	53	94.6	25	4	US-09-534-639-17
7	53	94.6	28	3	US-09-075-257A-15
8	53	94.6	28	3	US-09-075-257A-16
9	53	94.6	28	4	US-09-534-639-15
10	53	94.6	28	4	US-09-534-639-16
11	49	87.5	12	4	US-08-788-822A-21
12	33	58.9	13	1	US-08-305-871A-20
13	33	58.9	13	4	US-08-788-822A-24
14	33	58.9	1253	4	US-08-864-785-2
15	32	57.1	13	1	US-08-305-871A-21
16	32	57.1	13	4	US-08-788-822A-25
17	32	57.1	16	4	US-09-284-625-25
18	32	57.1	372	1	US-08-196-218-33
19	32	57.1	372	1	US-08-681-953-33
20	32	57.1	662	4	US-09-232-191-11
21	32	57.1	662	4	US-09-232-200-11
22	32	57.1	662	4	US-09-232-200-96
23	32	57.1	662	4	US-09-232-197-11
24	32	57.1	662	4	US-09-232-197-96
25	32	57.1	662	4	US-09-232-201-11
26	32	57.1	662	4	US-09-232-201-96
27	32	57.1	689	4	US-09-232-200-73

28 32 57.1 689 4 US-09-232-197-73  
29 32 57.1 689 4 US-09-232-201-73  
30 32 57.1 695 4 US-09-458-481B-4  
31 32 57.1 695 4 US-09-458-481B-5  
32 32 57.1 695 4 US-09-458-481B-6  
33 32 57.1 830 1 US-08-110-158-4  
34 32 57.1 830 5 PCT-US91-05059-2  
35 32 57.1 830 6 5378464-2  
36 31 55.4 297 3 US-09-058-489-10  
37 31 55.4 379 1 US-07-723-002C-6  
38 31 55.4 402 1 US-08-460-806-17  
39 31 55.4 402 1 US-08-325-630-17  
40 31 55.4 472 2 US-08-459-346-17  
41 31 55.4 472 3 US-08-889-419-17  
42 31 55.4 472 4 US-08-402-542-17  
43 31 55.4 472 5 PCT-US93-07189-17  
44 31 55.4 893 4 US-09-514-302-4  
45 31 55.4 1938 4 US-09-514-302-2

ALIGNMENTS

RESULT 1  
US-08-305-871A-22  
; Sequence 22, Application US/08305871A  
; Patent No. 5736142  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Gaeta, Federico  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sidney, John  
; APPLICANT: Alexander, Jeffrey L.  
; TITLE OF INVENTION: Alteration of Immune Response Using Pan  
; TITLE OF INVENTION: DR-Binding Peptides  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/305, 871A  
; FILING DATE: 14-SEP-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/121,101  
; FILING DATE: 14-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 14137-0062-10  
; TELEPHONE: (415) 576-0300  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..13  
; OTHER INFORMATION: /note= "Peptide wherein X is

OTHER INFORMATION: tyrosine or phenylalanine."  
US-08-305-871A-22

Query Match 94.6%; Score 53; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
|||||

DB 1 AKXVAAWTLKAA 12  
|||||

## RESULT 2

US-08-305-871A-23  
; Sequence 23, Application US/08305871A  
; Patent No. 5736142

GENERAL INFORMATION:

APPLICANT: Sette, Alessandro

APPLICANT: Gaeta, Federico

APPLICANT: Grey, Howard M.

APPLICANT: Sidney, John

APPLICANT: Alexander, Jeffrey L.

TITLE OF INVENTION: Alteration of Immune Response Using Pan

TITLE OF INVENTION: DR-Binding Peptides

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/305,871A

FILING DATE: 14-SEP-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/121,101

FILING DATE: 14-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 14137-0062-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1-13

OTHER INFORMATION: /note= "Peptide wherein X is

OTHER INFORMATION: tyrosine or phenylalanine."

## US-08-305-871A-23

Query Match 94.6%; Score 53; DB 1; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0027;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
|||||

DB 1 AKXVAAWTLKAA 12  
|||||

## RESULT 3

US-08-788-822A-26  
; Sequence 26, Application US/08788822A  
; Patent No. 6413935

GENERAL INFORMATION:

APPLICANT: Alexander, Jeffrey L.

APPLICANT: Defrees, Shawn

APPLICANT: Sette, Alessandro

TITLE OF INVENTION: Induction of Immune Response Against

TITLE OF INVENTION: Desired Determinants

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/788,822A

FILING DATE: 23-JAN-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/010,510

FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 014137-0092100S

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 3

OTHER INFORMATION: /product= "OTHER"

OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"

## US-08-788-822A-26

Query Match 94.6%; Score 53; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
|||||

DB 1 AKXVAAWTLKAA 12  
|||||

## RESULT 4

US-08-788-822A-27  
; Sequence 27, Application US/08788822A  
; Patent No. 6413935

GENERAL INFORMATION:

APPLICANT: Alexander, Jeffrey L.

APPLICANT: Defrees, Shawn

APPLICANT: Sette, Alessandro

TITLE OF INVENTION: Induction of Immune Response Against

TITLE OF INVENTION: Desired Determinants

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/788,822A  
FILING DATE: 23-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/010,510  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 014137-009210US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: peptide  
US-08-788-822A-27

Query Match 94.6%; Score 53; DB 4; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0027;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
||| |||||

Db 1 AKFVAAWTLKAA 12

## RESULT 5

US-09-075-257A-17  
Sequence 17, Application US/09075257A  
Patent No. 6074645  
GENERAL INFORMATION:  
APPLICANT: DIAMOND, DON JEFFREY  
APPLICANT: YORCK, JOANNE  
TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BART G. NEWLAND  
STREET: 555 13TH STREET, NW SUITE 701E  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,257A  
FILING DATE: 11-MAY-1998  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/021,298  
FILING DATE: 10-FEB-1998  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/950,064  
FILING DATE: 14-OCT-1997  
PRIOR APPLICATION DATA: US 08/747,488  
APPLICATION NUMBER: US 08/747,488  
FILING DATE: 12-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: NEWLAND, BART G  
REGISTRATION NUMBER: 31,282  
REFERENCE/DOCKET NUMBER: 1954-112CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 6  
OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
or phenylalanine"  
US-09-075-257A-17

Query Match 94.6%; Score 53; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
||| |||||

Db 4 AKXVAAWTLKAA 15

## RESULT 6

US-09-534-639-17  
Sequence 17, Application US/09534639  
Patent No. 6251399  
GENERAL INFORMATION:  
APPLICANT: Diamond, Don J  
APPLICANT: York, Joanne  
TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
TITLE OF INVENTION: CYTOMEGALOVIRUS  
FILE REFERENCE: 1954-343  
CURRENT APPLICATION NUMBER: US/09/534,639  
CURRENT FILING DATE: 2000-03-27  
EARLIER APPLICATION NUMBER: 09/075,257  
EARLIER FILING DATE: 1998-05-11  
EARLIER APPLICATION NUMBER: 09/021,298  
EARLIER FILING DATE: 1998-02-10  
EARLIER APPLICATION NUMBER: 08/950,064  
EARLIER FILING DATE: 1997-10-14  
EARLIER APPLICATION NUMBER: 08/747,488  
EARLIER FILING DATE: 1996-11-12  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (6)  
OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
OTHER INFORMATION: Description of Artificial Sequence: Human  
OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-17

Query Match 94.6%; Score 53; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.0052;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12  
Db 4 AKXVAWTLKAA 15

## RESULT 7

US-09-075-257A-15  
; Sequence 15, Application US/09075257A  
; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YORK, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPIOTOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 6  
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
; OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-15

Query Match 94.6%; Score 53; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1.AKXVAWTLKAA 12  
Db 4 AKXVAWTLKAA 15

## RESULT 8

US-09-075-257A-16  
; Sequence 16, Application US/09075257A  
; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YORK, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPIOTOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 6  
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
; OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-16

Query Match 94.6%; Score 53; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12  
Db 4 AKXVAWTLKAA 15

## RESULT 9

US-09-534-639-15  
; Sequence 15, Application US/09534639  
; Patent No. 6251399  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J  
; APPLICANT: York, Joanne

;; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
;; FILE REFERENCE: 1954-343  
;; CURRENT APPLICATION NUMBER: US/09/534,639  
;; EARLIER FILING DATE: 2000-03-27  
;; EARLIER APPLICATION NUMBER: 09/075,257  
;; EARLIER FILING DATE: 1998-05-11  
;; EARLIER APPLICATION NUMBER: 09/021,298  
;; EARLIER FILING DATE: 1998-02-10  
;; EARLIER APPLICATION NUMBER: 08/950,064  
;; EARLIER FILING DATE: 1997-10-14  
;; EARLIER APPLICATION NUMBER: 08/747,488  
;; EARLIER FILING DATE: 1996-11-12  
;; NUMBER OF SEQ ID NOS: 20  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 15  
;; LENGTH: 28  
;; TYPE: PPT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; NAME/KEY: VARIANT  
;; LOCATION: (6)  
;; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Human  
;; OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-15

Query Match 94.6%; Score 53; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKAA 12  
Db 4 AKXVAAWTLKAA 15

RESULT 10  
US-09-534-639-16  
;; Sequence 16, Application US/09534639  
;; Patent No. 6251399  
;; GENERAL INFORMATION:  
;; APPLICANT: Diamond, Don J  
;; APPLICANT: York, Joanne  
;; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
;; FILE REFERENCE: 1954-343  
;; CURRENT APPLICATION NUMBER: US/09/534,639  
;; CURRENT FILING DATE: 2000-03-27  
;; EARLIER APPLICATION NUMBER: 09/075,257  
;; EARLIER FILING DATE: 1998-05-11  
;; EARLIER APPLICATION NUMBER: 09/021,298  
;; EARLIER FILING DATE: 1998-02-10  
;; EARLIER APPLICATION NUMBER: 08/950,064  
;; EARLIER FILING DATE: 1997-10-14  
;; EARLIER APPLICATION NUMBER: 08/747,488  
;; EARLIER FILING DATE: 1996-11-12  
;; NUMBER OF SEQ ID NOS: 20  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 16  
;; LENGTH: 28  
;; TYPE: PPT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; NAME/KEY: VARIANT  
;; LOCATION: (6)  
;; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Human  
;; OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-16

Query Match 94.6%; Score 53; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKAA 12  
Db 4 AKXVAAWTLKAA 15

RESULT 11  
US-08-788-822A-21  
;; Sequence 21, Application US/08788822A  
;; Patent No. 6413935  
;; GENERAL INFORMATION:  
;; APPLICANT: Alexander, Jeffrey L.  
;; APPLICANT: DeFrees, Shawn  
;; APPLICANT: Sette, Alessandro  
;; TITLE OF INVENTION: Induction of Immune Response Against  
;; TITLE OF INVENTION: Desired Determinants  
;; NUMBER OF SEQUENCES: 30  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend and Crew LLP  
;; STREET: Two Embarcadero Center, Eighth Floor  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94111-3834  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/788,822A  
;; FILING DATE: 23-JAN-1997  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/010,510  
;; FILING DATE: 24-JAN-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Bastian, Kevin L.  
;; REGISTRATION NUMBER: 34,774  
;; REFERENCE/DOCKET NUMBER: 014137-009210US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 576-0200  
;; TELEFAX: (415) 576-0300  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 12 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 3  
;; OTHER INFORMATION: /product= "OTHER"  
;; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 12  
;; OTHER INFORMATION: /product= "OTHER"  
;; OTHER INFORMATION: /note= "Xaa = alaninamide"  
US-08-788-822A-21

Query Match 87.5%; Score 49; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKXVAAWTLKA 11  
Db 1 AKXVAAWTLKA 11

```
RESULT 12
US-08-305-871A-20
; Sequence 20, Application US/08305871A
; Patent No. 5736142
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Gaeta, Federico
; APPLICANT: Grey, Howard M.
; APPLICANT: Sidney, John
; APPLICANT: Alexander, Jeffrey L.
; TITLE OF INVENTION: Alteration of Immune Response Using Pan
; TITLE OF INVENTION: DR-Binding Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /notes= "Peptide wherein X is
; OTHER INFORMATION: tyrosine or phenylalanine."
US-08-305-871A-20
Query Match 58.9%; Score 33; DB 1; Length 13;
Best Local Similarity 83.3%; Pred. No. 6.6;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAATLKAA 12
Db 1 AAXVAATLKAA 12

RESULT 13
US-08-788-822A-24
; Sequence 24, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: Defrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"
US-08-788-822A-24
Query Match 58.9%; Score 33; DB 4; Length 13;
Best Local Similarity 83.3%; Pred. No. 6.6;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAATLKAA 12
Db 1 AAXVAATLKAA 12

RESULT 14
US-08-864-785-2
; Sequence 2, Application US/08864785A
; Patent No. 6329566
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Joshua M.
; APPLICANT: Oppenheimer, Allison J.
; APPLICANT: Hart, Anne C.
; TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,
; TITLE OF INVENTION: AND PREVENTION OF NEURODEGENERATION
; FILE REFERENCE: 00786/353001
; CURRENT APPLICATION NUMBER: US/08/864,785A
; CURRENT FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-08-864-785-2
Query Match 58.9%; Score 33; DB 4; Length 1253;
Best Local Similarity 58.3%; Pred. No. 6.4e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```



QY 1 AKXVAATLKA 12  
| : ||||| |  
Db 74 AVVIAATLYLA 85

RESULT 15  
US-08-305-871A-21  
; Sequence 21, Application US/08305871A  
; Patent No. 5736142  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Gaeta, Federico  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sidney, John  
; APPLICANT: Alexander, Jeffrey L.  
; TITLE OF INVENTION: Alteration of Immune Response Using Pan  
; TITLE OF INVENTION: DR-Binding Peptides  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/305.871A  
; FILING DATE: 14-SEP-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/121,101  
; FILING DATE: 14-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 14137-0062-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..13  
; OTHER INFORMATION: /note= "Peptide wherein X is  
; OTHER INFORMATION: tyrosine or phenylalanine."  
US-08-305-871A-21

Query Match 57.1%; Score 32; DB 1; Length 13;  
Best Local Similarity 75.0%; Pred. No. 9.8;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAATLKA 12  
| : ||||| |  
Db 1 MAXIAATLKA 12

Search completed: January 29, 2003, 10:48:21  
Job time : 10.4545 secs



GenCore version 5.1.3  
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:29:22 ; Search time 26.9455 Seconds  
(without alignments)  
64.288 Million cell updates/sec

Title: US-09-707-738-22

Perfect score: 56

Sequence: 1 AKXVAATLKAAX 13

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 13250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	94.6	13	16 AAR70249	Pan DR-binding pep
2	53	94.6	13	16 AAR70250	Pan DR-binding pep
3	53	94.6	13	18 AAR70251	Pan DR-binding pep
4	53	94.6	13	18 AAR70252	Padre (pan-DR bind
5	53	94.6	13	19 AAR70253	Padre (pan-DR bind
6	53	94.6	13	19 AAR70254	Pan DR binding pep
7	53	94.6	13	21 AAR70255	Pan DR binding pep
8	53	94.6	13	21 AAR70256	Promiscuous T help
9	53	94.6	13	21 AAR70257	Universal helper T
10	53	94.6	13	22 AAR70258	Pan-DR binding pep
	53	94.6	13	22 AAR70259	Pan-DR-binding pep

11	53	94.6	13	22 AAB99710	Pan-DR-binding pep
12	53	94.6	13	22 AAB99711	Pan-DR-binding pep
13	53	94.6	13	22 AAG84519	Pan-DR-binding epi
14	53	94.6	13	22 AAG88271	Pan-DR-binding epi
15	53	94.6	13	22 AAB20154	PADRE peptide. Sy
16	53	94.6	13	22 AAB46167	PADRE universal T
17	53	94.6	13	22 AAB49066	PADRE T-cell epitope
18	53	94.6	13	23 ABJ01954	158P1D7 related HL
19	53	94.6	13	23 AAU08293	Pan DR epitope pep
20	53	94.6	13	23 ABG34860	Pan-DR-binding epi
21	53	94.6	13	23 AAU91547	Pan-DR-binding epi
22	53	94.6	13	23 ABB94471	Pan-DR-binding epi
23	53	94.6	13	23 AAU10848	Helper CD4 peptide
24	53	94.6	15	18 AAW22123	Padre (pan-DR bind
25	53	94.6	20	22 AAB46179	Tetanus toxoid epi
26	53	94.6	20	22 AAB46182	Tetanus toxoid epi
27	53	94.6	20	22 AAB46204	Human APP A-beta 1
28	53	94.6	20	22 AAB49078	Amyloid beta/PADRE
29	53	94.6	20	22 AAB49081	Amyloid beta/PADRE
30	53	94.6	22	23 AAU10836	Human cytomagalovi
31	53	94.6	22	23 AAU10839	Human cytomagalovi
32	53	94.6	22	23 AAU10842	Human cytomagalovi
33	53	94.6	22	23 AAU10843	Human cytomagalovi
34	53	94.6	23	23 AAU10834	Human cytomagalovi
35	53	94.6	24	23 ABB79184	Human cytomagalovi
36	53	94.6	25	20 AAU09330	Molecule A SEQ ID
37	53	94.6	25	21 AAB12422	Human cytomagalovi
38	53	94.6	25	22 AAE05441	Human cytomagalovi
39	53	94.6	25	23 AAU10844	Human cytomagalovi
40	53	94.6	25	23 AAU10854	Human cytomagalovi
41	53	94.6	27	23 ABB79178	Human cytomagalovi
42	53	94.6	27	23 ABB79181	Human cytomagalovi
43	53	94.6	28	19 AAW48993	Lipidated vaccine
44	53	94.6	28	19 AAW48994	Lipidated vaccine
45	53	94.6	28	20 AAU09328	Lipidated vaccine

#### ALIGNMENTS

RESULT 1  
AAR70249  
ID AAR70249 standard; peptide; 13 AA.  
XX  
AC AAR70249;  
XX  
DT 13-NOV-1995 (first entry)  
XX  
DE Pan DR-binding peptide activator of T cells.  
XX  
KW T cell; MHC; class II; major histocompatibility complex II;  
KW immunogen; activation.  
XX  
OS Synthetic.  
XX  
PH Key  
FT Misc-difference 1 Location/Qualifiers  
FT Modified-site 3 /note= "D-form residue"  
FT FT /label= OTHER  
FT FT /note= "cyclohexylalanine"  
FT FT Misc-difference 13  
FT FT /note= "D-form residue"

XX WO9507707-A.

XX 23-MAR-1995.

XX 14-SEP-1994; 94WO-US10368.

XX 14-SEP-1993; 93US-0121101.

XX (CYTE-) CYTEL CORP.

XX Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
 XX WPI; 1995-131178/17.  
 XX Inhibiting or inducing an immune response using Pan DR-binding  
 PT peptide(s) - to prepare a compsn. used to treat allo:graft  
 PT rejection, allergic response and auto:immunity and as a vaccine  
 PT component  
 XX Claim 3; Page 51; 59pp; English.  
 XX Peptides capable of binding MHC class II molecules via antigen  
 CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
 CC AAR70249 is capable of binding MHC class II molecules causing T cell  
 CC activation. This is useful as a vaccine component and may be used in  
 CC conjunction with CTL peptides to enhance a CTL response.  
 XX Sequence 13 AA;  
 SQ Query Match 94.6%; Score 53; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.0058;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKXVAAWTLKAA 12  
 Db |||||  
 1 AKXVAAWTLKAA 12  
 RESULT 2  
 AAR70250 standard; peptide; 13 AA.  
 XX AC AAR70250;  
 XX DT 13-NOV-1995 (first entry)  
 XX DE Pan DR-binding peptide activator of T cells.  
 XX KW T cell; MHC; class II; major histocompatibility complex II;  
 XX immunogen; activation.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT W09507707-A.  
 XX DT 23-MAR-1995.  
 XX PF 14-SEP-1994; 94WO-US10368.  
 XX PR 14-SEP-1993; 93US-0121101.  
 XX PA (CYTE-) CYTEL CORP.  
 XX PI Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
 XX WPI; 1995-131178/17.  
 XX Inhibiting or inducing an immune response using Pan DR-binding  
 PT peptide(s) - to prepare a compsn. used to treat allo:graft  
 PT rejection, allergic response and auto:immunity and as a vaccine  
 PT component  
 XX Claim 3; Page 51; 59pp; English.  
 XX Peptides capable of binding MHC class II molecules via antigen  
 CC binding sites (pan DR-binding peptides) were synthesised. Peptide

CC AAR70250 is capable of binding MHC class II molecules causing T cell  
 CC activation. This can be used as a vaccine component. The peptide can also  
 CC be used in combination with CTL peptides to enhance a CTL response.  
 XX Sequence 13 AA;  
 SQ Query Match 94.6%; Score 53; DB 16; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.0059;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKXVAAWTLKAA 12  
 Db |||||  
 1 AKXVAAWTLKAA 12  
 RESULT 3  
 AAW22120 standard; peptide; 13 AA.  
 XX AC AAW22120;  
 XX DT 13-MAR-1998 (first entry)  
 XX DE Padre (pan-DR binding) peptide 5.  
 XX KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Modified-site 3 /note= "cyclohexylalanine"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT W09726784-A1.  
 XX DT 31-JUL-1997.  
 XX PF 23-JAN-1997; 97WO-US01041.  
 XX PR 24-JAN-1996; 96US-0010510.  
 XX PA (CYTE-) CYTEL CORP.  
 XX PI Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 XX Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 PT prevention and treatment of tumours and infections  
 XX Claim 23; Page 74; 87pp; English.  
 XX This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction

CC of cytotoxic T cells and antibody production.

SQ Sequence 13 AA;

Query Match 94.6%; Score 53; DB 18; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.0058;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12

DB 1 AKXVAAWTLKAA 12

#### RESULT 4

AAW22121  
ID AAW22121 standard; peptide; 13 AA.

XX

AC AAW22121;

XX 13-MAR-1998 (first entry)

XX Padre (pan-DR binding) peptide 6.

DE

XX PADRE; pan-DR binding; immune response; antigenic determinant;  
treatment; tumour; infection.

XX Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 13 /note= "D-form residue"

XX WO9726784-A1.

XX 31-JUL-1997.

XX

XX 23-JAN-1997; 97WO-US01041.

XX 24-JAN-1996; 96US-0010510.

XX (CYTE-) CYTEL CORP.

XX Alexander JL, Defrees S, Sette A;

XX WPI; 1997-393272/36.

XX

PT Composition for eliciting immune response to non-protein determinant  
- comprises the determinant and a pan-DR binding peptide, used for  
prevention and treatment of tumours and infections

XX Claim 23; Page 74; 87pp; English.

XX This PADRE (pan-DR binding) peptide is used in a composition for  
eliciting an immune response to a non-protein antigenic determinant.

XX The composition comprises of the PADRE peptide covalently linked to the  
non-protein antigenic determinant. The composition is used to induce a  
therapeutic or prophylactic response, particularly to selected  
polysaccharide antigens associated with tumours or infectious agents. It  
provides a high level, long-lasting IGC immune response. The composition  
is also used to produce monoclonal antibodies which are potentially  
useful as therapeutic and diagnostic agents. The composition can be used  
to diagnose susceptibility of a patient to treatment with the non-protein  
antigenic determinant or to predict subjects at risk from developing  
chronic infections. PADRE peptides have broader specificity and higher  
affinity than known DR-binding peptides. They are powerful inhibitors of  
the proliferative response of human T cells restricted by at least 6  
different DR molecules, and act as helper epitopes of in vivo induction  
of cytotoxic T cells and antibody production.

XX SQ Sequence 13 AA;

Query Match 94.6%; Score 53; DB 18; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.0058;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12

DB 1 AKXVAAWTLKAA 12

#### RESULT 5

AAW50125  
ID AAW50125 standard; peptide; 13 AA.

XX

AC AAW50125;

XX 30-JUN-1998 (first entry)

XX Pan DR binding peptide (22).

XX

XX Pan DR binding peptide; antigen binding site; MHC molecule;

XX DR locus.

XX Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 3 /label= Tyr, Phe

XX US5736142-A.

XX 07-APR-1998.

XX 14-SEP-1994; 94US-0305871.

XX 14-SEP-1994; 94US-0305871.

XX 14-SEP-1993; 93US-0121101.

XX (CYTE-) CYTEL CORP.

XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;

XX WPI; 1998-239154/21.

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
or inducing MHC Class II mediated activation of T cells

XX Claim 35; Columns 39-40; 29pp; English.

XX The present sequence, a pan DR binding peptide, is capable of  
binding antigen binding sites on MHC molecules, which are encoded  
by most of the alleles of a DR locus. The peptide can be used to  
inhibit or induce MHC Class II mediated activation of T-cells or  
helper T-cells, which themselves mediate a CTL response. The  
peptide can be used in mammals, especially humans, to inhibit  
T-cell-mediated events involved in allograft rejection, allergic  
responses and autoimmunity and as a vaccine adjuvant for enhancing  
an immune response against an administered immunogen. The peptide  
can be used with other immunogens to treat, e.g. prostate cancer,  
hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
lymphoma, CMV and condyloma acuminatum.

XX SQ Sequence 13 AA;

Query Match 94.6%; Score 53; DB 19; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.0058;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12

DB 1 AKXVAAWTLKAA 12

#### RESULT 6

AAW50126  
ID AAW50126 standard; peptide; 13 AA.

AC AAW50126;

DT 30-JUN-1998 (first entry)

DE Pan DR binding peptide (23).

KW Pan DR binding peptide; antigen binding site; MHC molecule;  
DR locus.

KW Synthetic.

OS  
FH Key Location/Qualifiers  
FT Misc-difference 3 /label= Tyr, Phe

XX US5736142-A.

XX 07-APR-1998.

XX 14-SEP-1994; 94US-0305871.

XX 14-SEP-1994; 94US-0305871.

XX 14-SEP-1993; 93US-0121101.

XX (CYTE-) CYTEL CORP.

XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;

XX WPI; 1998-239154/21.

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting

PT or inducing MHC Class II mediated activation of T cells

PS Claim 35; Columns 41-42; 29pp; English.

XX The present sequence, a pan DR binding peptide, is capable of  
CC binding antigen binding sites on MHC molecules, which are encoded  
CC by most of the alleles of a DR locus. The peptide can be used to  
CC inhibit or induce MHC Class II mediated activation of T-cells or  
CC helper T-cells, which themselves mediate a CTL response. The  
CC peptide can be used in mammals, especially humans, to inhibit  
CC T-cell-mediated events involved in allograft rejection, allergic  
CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
CC an immune response against an administered immunogen. The peptide  
CC can be used with other immunogens to treat, e.g. prostate cancer,  
CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
CC lymphoma, CMV and condyloma acuminatum.

XX Sequence 13 AA;

Query Match 94.6%; Score 53; DB 19; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
Db 1 AKXVAAWTLKAA 12

RESULT 7

ID AAB36289 standard; Peptide; 13 AA.

AC AAB36289;

XX 26-FEB-2001 (first entry)

DE Promiscuous T helper epitope SEQ ID NO: 65.

KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;

KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX Synthetic.

XX WO200065058-A1.

XX 02-NOV-2000.

XX 19-APR-2000; 2000WO-DK00205.

XX 23-APR-1999; 99DK-0000552.

XX 06-MAY-1999; 99US-0132811.

XX (MBBI-) M & E BIOTECH AS.

XX Klysner S;

XX WPI; 2000-672791/65.

XX Down-regulating interleukin 5 (IL-5) activity in humans by  
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,  
PT prophylaxis or amelioration of asthma or other chronic allergic  
PT conditions -  
XX Disclosure; Page 168; 172pp; English.

XX The present invention is concerned with methods of treating asthma,  
CC eosinophilia, allergic rhinitis and other allergic diseases. These  
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
CC proteins and their coding sequences to down-regulate IL-5 activity and  
CC thus reduce eosinophil numbers. The allergic diseases may be treated  
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
CC it is possible that they may be used in the treatment of cancer and  
CC helminthic infections.

XX Sequence 13 AA;

Query Match 94.6%; Score 53; DB 21; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0058;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
Db 1 AKXVAAWTLKAA 12

RESULT 8

ID AAY52558 standard; peptide; 13 AA.

XX AAY52558;

XX 28-FEB-2000 (first entry)

XX Universal helper T epitope, pan DR epitope (PADRE).

XX Chimeric; Ii protein; pan DR epitope; expression vector;  
KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
KW epitope; antigen; presentation; class I; cytosolic pathway;  
KW endoplasmic reticulum; class II; extracellular antigen;  
KW endocytic pathway; helper T lymphocyte; HII; universal epitope;  
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
KW autoimmune disease; activation; antiviral; antimalarial;  
KW immunoprotective.

XX Synthetic.

XX WO9958658-A2.

XX 18-NOV-1999.

PF 13-MAY-1999; 99WO-US10646.  
 XX 13-MAY-1998; 98US-0078904.  
 PR 15-MAY-1998; 98US-0085751.  
 XX (EPIM-) EPIMMUNE INC.  
 XX Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
 PI Chesnut RW;  
 XX WPI; 2000-039103/03.  
 DR N-PSDB; AAZ38679.  
 XX Expression vectors encoding major histocompatibility targeting  
 PT sequence, used as, e.g. tumor vaccines -  
 XX Claim 9; Page 80; 130pp; English.  
 CC This sequence represents a universal helper T epitope, pan DR epitope  
 CC (PADRE), DNA encoding which is used to construct fusion genes used  
 CC in exemplifications of the present invention. The invention  
 CC relates to a novel expression vector comprising a promoter operably  
 CC linked to a fusion gene encoding a major histocompatibility complex  
 CC (MHC) targeting sequence, and two or more heterologous peptide epitopes.  
 CC The MHC targeting sequence may be a class I targeting sequence, which  
 CC directs an MHC class I epitope to a cytosolic pathway or to the  
 CC endoplasmic reticulum, or an MHC class II targeting sequence, which  
 CC directs extracellular antigens to enter the endocytic pathway to be  
 CC processed into antigen peptides for presentation on MHC class II  
 CC molecules. The heterologous epitopes may comprise either helper T  
 CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and  
 CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are  
 CC useful for stimulating an immune response in vivo, as well as for use in  
 CC assaying the human immunogenicity of a human T cell peptide epitope in  
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
 CC enhancing immunity against infectious pathogens such as viruses (e.g.,  
 CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
 CC Plasmodium falciparum, the cause of malaria) and also tumor cells and  
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
 CC combined with other MHC class I and class II epitopes to increase the  
 CC number of cells that are activated in response to a given antigen and  
 CC provide a broader population coverage of MHC-reactive alleles.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 94.6%; Score 53; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.0058;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 1 AKFVAAWTLKAA 12  
 |||||  
 RESULT 9  
 AAB73644  
 ID AAB73644 standard; peptide; 13 AA.  
 XX  
 AC AAB73644;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Pan-DR binding peptide, PADRE.  
 XX  
 KW Cellular vaccine; antigen-presenting cell; APC;  
 KW cell-surface molecule density; major histocompatibility complex; MHC;  
 KW antigen-specific T-cell expansion; tumour; cancer; viral infection;  
 KW parasitic infection; Pan-DR binding peptide; PADRE.  
 XX  
 OS Unidentified.  
 XX  
 PN W0200136978-A1.  
 XX

PD 25-MAY-2001.  
 XX  
 PF 15-NOV-2000; 2000WO-US42213.  
 XX  
 PR 15-NOV-1999; 99US-0165428.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schultze JL, Vonderheide RH, Nadler LM, Maecker B;  
 PI Von Bergwelt-Baildon M;  
 XX WPI; 2001-343909/36.  
 DR  
 XX Determining surface density of molecules on antigen-presenting cells,  
 PT useful for quality control of cellular vaccines for treatment of tumors  
 PT and infections comprises measuring density ratios -  
 XX  
 PS Disclosure; Page 21; 47pp; English.  
 XX  
 CC The invention relates to a method of determining the surface density of  
 CC a cell-surface molecule on a primary or artificial antigen-presenting  
 CC cell (APC). The method comprises determining the cell surface area  
 CC of the APCs, determining the absolute amount of cell surface molecule on  
 CC the surface of the APCs, and calculating the ratio of the amount of  
 CC the cell surface molecule to the APC surface area as a measure of cell  
 CC surface molecule density. The invention also encompasses a kit for  
 CC determining the cell-surface density of a complex of peptide and MHC  
 CC (major histocompatibility complex) protein/peptide complex on an APC;  
 CC identifying an agent that increases persistence of the MHC/peptide  
 CC complex on an APC surface by culturing APCs in presence of test compound  
 CC and measuring the time of persistence relative to an untreated control;  
 CC and determining whether the APC carries a therapeutically adequate amount of  
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The  
 CC method is used to determine if the amount of MHC/antigenic peptide  
 CC complex present on the surface of an APC is sufficient for therapeutic  
 CC use of the cells as cellular vaccines or for ex vivo expansion of  
 CC antigen-specific T cells for subsequent return to the patient.  
 CC particularly for the treatment of tumours or viral or parasitic  
 CC infections. The method can also be used to identify agents that increase  
 CC production of MHC/peptide complexes on cells, (which increases the  
 CC therapeutic potential of the cells) by treatment before and/or during  
 CC APC administration. The method allows identification of therapeutically  
 CC useful antigen-pulsed APC optimisation of conditions for their  
 CC production, and control of APC quality. Sequences AAB73643-AAB73648  
 CC represent peptides used in MHC binding studies and for the generation  
 CC of peptide-specific cytotoxic T-lymphocytes.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 94.6%; Score 53; DB 22; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.0058;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKXVAAWTLKAA 12  
 |||||  
 Db 1 AKFVAAWTLKAA 12  
 |||||  
 RESULT 10  
 AAB99709  
 ID AAB99709 standard; peptide; 13 AA.  
 XX  
 AC AAB99709;  
 XX  
 DT 06-SEP-2001 (first entry)  
 XX  
 DE Pan-DR-binding peptide (PADRE) SEQ ID NO:30.  
 XX  
 KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;  
 KW cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;  
 KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;  
 KW immunotherapy; immune response.

XX Homo sapiens.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Modified-site 3 /note= "cyclohexylalanine"  
XX

PN WO200141741-A1.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-US34318.

PR 13-DEC-1999; 99US-0170448.

PR 05-APR-2000; 2000US-0543608.

PR 30-MAY-2000; 2000US-0583200.

XX (EPIM-) EPIMMUNE INC.

XX Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;

PI Chesnut R;

XX WPI; 2001-381489/40.

XX Compositions for use in a vaccine for treating, e.g., breast, lung and  
PT colon cancer comprises at least one peptide that comprises an isolated  
PT epitope of a tumor-associated antigen -

XX Example 7; Page 47; 86pp; English.

XX The present invention describes a composition (I) comprising at least  
CC one peptide that comprises an isolated, prepared epitope consisting of  
CC a sequence selected from 25 short amino acid sequences given in AAB99680  
CC to AAB99704. Also described are: (1) a composition (II) comprising one  
CC or more peptides, and further comprising at least two epitopes selected  
CC from the 25 short amino acid sequences (as above), where each of the one  
CC or more peptides comprise less than 50 contiguous amino acids that have  
CC 100% identity with a native peptide sequence; and (2) a vaccine  
CC composition (III) comprising an epitope selected from the 25 short amino  
CC acid sequences (as above) and a pharmaceutical excipient. (I) has  
CC cytostatic and immunomodulatory activities and can be used in vaccine  
CC production and immunotherapy. The peptide epitope compositions (I)-(II)  
CC are useful for monitoring an immune response to a tumour associated  
CC antigen or when one or more peptides are combined to create a vaccine  
CC (III) that stimulates the cellular arm of the immune system. In  
CC particular, the vaccine mediates immune responses against tumours in  
CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2  
CC supertype and improve the standard of care for patients being treated  
CC for breast, colon, or lung cancer. The present sequence represents a  
CC pan-DR-binding peptide (PADRE) sequence, which is used in an example  
CC from the present invention.

XX Sequence 13 AA;

Query Match 94.6%; Score 53; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12

Db 1 AKXVAAWTLKAA 12

RESULT 11

AAB99710

ID AAB99710 standard; peptide; 13 AA.

XX AAB99710;

XX 06-SEP-2001 (first entry)

DE Pan-DR-binding peptide (PADRE) SEQ ID NO:31.

XX Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;  
KW cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;  
KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;  
KW immunotherapy; immune response.

XX Homo sapiens.

OS Synthetic.

PN WO200141741-A1.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-US34318.

PR 13-DEC-1999; 99US-0170448.

PR 05-APR-2000; 2000US-0543608.

PR 30-MAY-2000; 2000US-0583200.

XX (EPIM-) EPIMMUNE INC.

XX Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;

PI Chesnut R;

XX WPI; 2001-381489/40.

XX Compositions for use in a vaccine for treating, e.g., breast, lung and  
PT colon cancer comprises at least one peptide that comprises an isolated  
PT epitope of a tumor-associated antigen -

XX Example 7; Page 48; 86pp; English.

XX The present invention describes a composition (I) comprising at least  
CC one peptide that comprises an isolated, prepared epitope consisting of  
CC a sequence selected from 25 short amino acid sequences given in AAB99680  
CC to AAB99704. Also described are: (1) a composition (II) comprising one  
CC or more peptides, and further comprising at least two epitopes selected  
CC from the 25 short amino acid sequences (as above), where each of the one  
CC or more peptides comprise less than 50 contiguous amino acids that have  
CC 100% identity with a native peptide sequence; and (2) a vaccine  
CC composition (III) comprising an epitope selected from the 25 short amino  
CC acid sequences (as above) and a pharmaceutical excipient. (I) has  
CC cytostatic and immunomodulatory activities and can be used in vaccine  
CC production and immunotherapy. The peptide epitope compositions (I)-(II)  
CC are useful for monitoring an immune response to a tumour associated  
CC antigen or when one or more peptides are combined to create a vaccine  
CC (III) that stimulates the cellular arm of the immune system. In  
CC particular, the vaccine mediates immune responses against tumours in  
CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2  
CC supertype and improve the standard of care for patients being treated  
CC for breast, colon, or lung cancer. The present sequence represents a  
CC pan-DR-binding peptide (PADRE) sequence, which is used in an example  
CC from the present invention.

XX Sequence 13 AA;

Query Match 94.6%; Score 53; DB 22; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0058;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12

Db 1 AKFVAAWTLKAA 12

RESULT 12

AAB99711

ID AAB99711 standard; peptide; 13 AA.

XX AAB99711;

XX 06-SEP-2001 (first entry)



DE XX Pan-DR-binding peptide (PADRE) SEQ ID NO:32.  
 KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;  
 KW Cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;  
 KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;  
 KW immunotherapy; immune response.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX XX WO200141741-A1.  
 PN 14-JUN-2001.  
 PD XX  
 XX 13-DEC-2000; 2000WO-US34318.  
 PF 13-DEC-1999; 99US-0170448.  
 XX 05-APR-2000; 2000US-0543608.  
 PR 30-MAY-2000; 2000US-0583200.  
 XX (EPIM-) EPIMUNE INC.  
 PA Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;  
 XX Chesnut R;  
 PI WPI; 2001-381489/40.  
 XX  
 DR Compositions for use in a vaccine for treating, e.g., breast, lung and  
 PT colon cancer comprises at least one peptide that comprises an isolated  
 PT epitope of a tumor-associated antigen -  
 XX  
 PS Example 7; Page 48; 86pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising at least  
 CC one peptide that comprises an isolated, prepared epitope consisting of  
 CC a sequence selected from 25 short amino acid sequences given in AAB99680  
 CC to AAB99704. Also described are: (1) a composition (II) comprising one  
 CC or more peptides, and further comprising at least two epitopes selected  
 CC from the 25 short amino acid sequences (as above), where each of the one  
 CC or more peptides comprise less than 50 contiguous amino acids that have  
 CC 100% identity with a native peptide sequence; and (2) a vaccine  
 CC composition (III) comprising an epitope selected from the 25 short amino  
 CC acid sequences (as above) and a pharmaceutical excipient. (I) has  
 CC cytostatic and immunomodulatory activities and can be used in vaccine  
 CC production and immunotherapy. The peptide epitope compositions (I)-(II)  
 CC are useful for monitoring an immune response to a tumour associated  
 CC antigen or when one or more peptides are combined to create a vaccine  
 CC (III) that stimulates the cellular arm of the immune system. In  
 CC particular, the vaccine mediates immune responses against tumours in  
 CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2  
 CC supertype and improve the standard of care for patients being treated  
 CC for breast, colon, or lung cancer. The present sequence represents a  
 CC pan-DR-binding peptide (PADRE) sequence, which is used in an example  
 CC from the present invention.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 94.6%; Score 53; DB 22; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.0058;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKYVAAWTLKAA 12  
 Db 1 AKYVAAWTLKAA 12  
 |||||||  
 RESULT 13  
 AAG84519  
 ID AAG84519 standard; Peptide; 13 AA.  
 XX  
 AC AAG84519;  
 XX  
 DT 10-SEP-2001 (first entry)

XX Pan-DR-binding epitope (PADRE) generic peptide.  
 DE Human; human leukocyte antigen; HLA epitope; cytotoxic T lymphocyte;  
 XX CTL; MAGE2; MAGE3; melanoma antigen gene; immune response; vaccine;  
 KW cancer; cytostatic; immunostimulant.  
 KW  
 XX OS Synthetic.  
 XX WO200142267-A1.  
 PN 14-JUN-2001.  
 PD XX  
 XX 11-DEC-2000; 2000WO-US33545.  
 PF 10-DEC-1999; 99US-0458298.  
 XX (EPIM-) EPIMUNE INC.  
 PA Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;  
 XX Keogh E;  
 PI WPI; 2001-375002/39.  
 DR  
 XX An isolated prepared MAGE2/3 epitope (I) for use in pharmaceuticals for  
 PT the treatment and prevention of cancer -  
 PT  
 PS Disclosure; Page 37; 171pp; English.  
 XX  
 CC The present invention describes MAGE2/3 epitopes (I). Also described  
 CC are: (1) a clonal cytotoxic T lymphocyte (CTL) that is cultured in vivo  
 CC and binds to a complex of (I); (2) a peptide (II) comprising (I) and a  
 CC second epitope and has less than 50 contiguous amino acids; (3) a vaccine  
 CC composition comprising (II), a unit dose of a peptide with at least 50  
 CC contiguous amino acids with 100% identity to the native peptide sequence  
 CC of MAGE2/3, and a pharmaceutical excipient; (4) an isolated nucleic acid  
 CC encoding (I); and (5) an isolated nucleic acid encoding (II). (I) has  
 CC cytostatic activity, and can be used in vaccines and as an  
 CC immunostimulant. A vaccine of (3) is useful for the treatment and  
 CC prevention of cancer. (I) is useful for monitoring or evaluating an  
 CC immune response by incubating a T-lymphocyte sample from a patient with  
 CC (I) that binds to an human leukocyte antigen (HLA) allele present in the  
 CC patient and detecting the presence of the T-lymphocyte that binds to the  
 CC peptide. The vaccine allows the opportunity to combine epitopes derived  
 CC from multiple tumour-associated molecules reducing the likelihood of  
 CC tumour escape due to antigen loss. AAG84515 to AAG84909 and AAB99725  
 CC represent amino acid sequences used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 94.6%; Score 53; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.0058;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKXVAAWTLKAA 12  
 Db 1 AKXVAAWTLKAA 12  
 |||||||  
 RESULT 14  
 AAG88271  
 ID AAG88271 standard; Peptide; 13 AA.  
 XX  
 AC AAG88271;  
 XX  
 DT 11-SEP-2001 (first entry)  
 DE  
 DE Pan-DR-binding epitope (PADRE) generic peptide.  
 KW Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;  
 KW immune response; vaccine; tumour; cancer; cytostatic; immunostimulant;  
 KW tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.

XX OS Synthetic.  
XX PN WO200141787-A1.  
XX PD 14-JUN-2001.  
XX PF 11-DEC-2000; 2000WO-US33591.  
XX PR 10-DEC-1999; 99US-0458299.  
XX PA (EPIM-) EPIMMUNE INC.  
XX PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;  
XX PI Keogh E;  
XX PI WPI; 2001-3749995/39.  
XX DR  
XX PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing  
XX PT cellular immune responses for the prevention and treatment of cancer -  
XX PS Disclosure; Page 37; 199pp; English.  
XX CC The present invention describes isolated prepared HER2/neu epitopes (I).  
XX CC Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is  
XX CC cultured in vitro and binds to a complex of an epitope (I), bound to a  
XX CC human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)  
XX CC and a second epitope and the peptide is less than 50 contiguous amino  
XX CC acids that have 100% identity with a native peptide sequence of HER2/neu;  
XX CC (3) a vaccine composition (III) comprising (II) and a pharmaceutical  
XX CC excipient; (4) an isolated nucleic acid encoding a peptide comprising  
XX CC (I); and (5) an isolated nucleic acid encoding (II). (I) has cytostatic  
XX CC and immunostimulant activities, and can be used in vaccines (I), (II)  
XX CC and (III) are useful for inducing cellular immune responses for the  
XX CC prevention and treatment of cancer. (I) and (II) are useful for  
XX CC monitoring or evaluating an immune response to a tumour-associated  
XX CC antigen when incubated with a T lymphocyte sample from a patient and  
XX CC detecting the presence of bound T lymphocyte to (I) or (II). Epitope  
XX CC based vaccines mean that immunosuppressive epitopes that may be present  
XX CC in whole antigens may be avoided. Selected epitopes may be combined to  
XX CC enhance immunogenicity. The possible pathological side effects caused by  
XX CC infectious agents or whole protein antigen is eliminated. The vaccine  
XX CC provides the ability to direct and focus an immune response to multiple  
XX CC selected antigens from the same pathogen. Epitope-based anti-tumour  
XX CC vaccines provides the opportunity to combine epitopes derived from  
XX CC multiple tumour-associated molecules addressing the problem of tumour-  
XX CC tumour variability and reducing the likelihood of tumour escape due to  
XX CC antigen loss. AAG89266 to AAG89121 represent amino acid sequences used in  
XX CC the exemplification of the present invention.  
XX SQ  
XX Sequence 13 AA;  
Query Match 94.6%; Score 53; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKXVAAWTLKAA 12  
Db 1 AKXVAAWTLKAA 12  
RESULT 15  
AAB20154  
ID AAB20154 standard; Peptide; 13 AA.  
XX AC AAB20154;  
XX DT 30-APR-2001 (first entry)  
XX DE PADRE peptide.  
XX KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;  
XX KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;

KW cardiant; PADRE; pan DE epitope.  
XX OS Synthetic.  
XX PN WO200105820-A2.  
XX PD 25-JAN-2001.  
XX PF 20-JUL-2000; 2000WO-DK00413.  
XX PR 20-JUL-1999; 99DK-0001014.  
XX PR 26-JUL-1999; 99US-0145275.  
XX PA (MEBI-) M & E BIOTECH AS.  
XX PI Halkier T, Mouritsen S, Klysner S;  
XX PI WPI; 2001-112680/12.  
XX DR  
XX PT Increasing the muscle mass of animals used in meat production by down  
XX PT regulating growth differentiation factor 8 (GDF-8) activity in the  
XX PT animal through induction of anti-GDF-8 antibody production -  
XX PS Disclosure; Page 15; 110pp; English.  
XX CC The present sequence is that of a PADRE (pan DE epitope) peptide  
XX CC which acts as a T-cell epitope and is capable of binding a large  
XX CC proportion of major histocompatibility complex class II molecules.  
XX CC It is an object of the invention to produce a recombinant  
XX CC therapeutic vaccine that is capable of effecting down-regulation of  
XX CC growth differentiation factor 8 (GDF-8) in order to increase the  
XX CC muscle growth rate of farm animals. The vaccines (see AAB20145-53)  
XX CC are capable of breaking autotolerance against autologous GDF-8.  
XX CC They comprise the C-terminal portion of human GDF-8 in which a  
XX CC portion of the native sequence is replaced by a T-cell epitope such  
XX CC as PADRE, with minimal disturbance of the authentic 3-dimensional  
XX CC structure of the protein. Down-regulation of GDF-8 activity can  
XX CC increase muscle mass by up to at least 45% in cattle, pigs and  
XX CC poultry used for meat production, reducing the need for antibiotic  
XX CC feed-additives. Anti-GDF8 vaccines can be used to treat human  
XX CC diseases such as cancer cachexia where muscle atrophy is pronounced  
XX CC and for patients suffering from acute and chronic heart failure.  
XX SQ  
XX Sequence 13 AA;  
Query Match 94.6%; Score 53; DB 22; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.0058;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12  
Db 1 AKXVAAWTLKAA 12

Search completed: January 29, 2003, 10:44:34  
Job time : 27.9455 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:37 ; Search time 10.9091 Seconds  
(without alignments)  
40.457 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXWTLKXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/6CTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	69.4	4	1	US-08-305-871A-17
2	25	69.4	4	4	US-08-788-822A-1
3	25	69.4	12	4	US-08-788-822A-21
4	25	69.4	13	1	US-08-305-871A-22
5	25	69.4	13	1	US-08-305-871A-23
6	25	69.4	13	4	US-08-788-822A-26
7	25	69.4	13	4	US-08-788-822A-27
8	25	69.4	14	4	US-08-788-822A-30
9	25	69.4	18	2	US-08-652-369A-1
10	25	69.4	23	4	US-07-741-453A-25
11	25	69.4	25	3	US-09-075-257A-17
12	25	69.4	25	4	US-09-534-639-17
13	25	69.4	28	3	US-09-075-257A-15
14	25	69.4	28	3	US-09-075-257A-16
15	25	69.4	28	4	US-09-534-639-15
16	25	69.4	28	4	US-09-534-639-16
17	25	69.4	86	4	US-08-858-207A-471
18	25	69.4	90	1	US-07-956-700B-85
19	25	69.4	90	1	US-08-476-537-85
20	25	69.4	90	1	US-08-485-607-85
21	25	69.4	90	2	US-08-475-879-85
22	25	69.4	90	4	US-09-433-043B-85
23	25	69.4	97	2	US-08-585-585A-7
24	25	69.4	97	2	US-08-249-037C-7
25	25	69.4	97	2	US-08-788-622B-7
26	25	69.4	97	3	US-08-788-621B-7
27	25	69.4	152	4	US-08-936-165A-264

28	25	69.4	169	4	US-09-134-001C-2948	Sequence 2948, Ap
29	25	69.4	170	4	US-09-134-001C-3512	Sequence 3512, Ap
30	25	69.4	173	4	US-09-134-001C-2944	Sequence 2944, Ap
31	25	69.4	221	4	US-09-247-373B-54	Sequence 54, Appl
32	25	69.4	225	4	US-09-071-035-376	Sequence 376, Appl
33	25	69.4	252	4	US-09-071-035-374	Sequence 374, Appl
34	25	69.4	277	4	US-07-741-453A-58	Sequence 58, Appl
35	25	69.4	352	1	US-08-482-577B-2	Sequence 2, Appli
36	25	69.4	352	3	US-08-289-222E-4	Sequence 4, Appli
37	25	69.4	352	4	US-09-218-176-2	Sequence 2, Appli
38	25	69.4	352	4	US-09-054-526B-4	Sequence 4, Appli
39	25	69.4	365	1	US-08-447-500-4	Sequence 4, Appli
40	25	69.4	365	1	US-08-454-097-4	Sequence 4, Appli
41	25	69.4	365	1	US-08-453-866-4	Sequence 4, Appli
42	25	69.4	365	3	US-09-100-664A-13	Sequence 13, Appl
43	25	69.4	365	3	US-08-185-359-4	Sequence 4, Appli
44	25	69.4	365	4	US-09-335-983-13	Sequence 13, Appl
45	25	69.4	391	4	US-07-741-453A-2	Sequence 2, Appli

## ALIGNMENTS

### RESULT 1

US-08-305-871A-17  
; Sequence 17, Application US/08305871A  
; Patent No. 5736142

### GENERAL INFORMATION:

APPLICANT: Sette, Alessandro  
APPLICANT: Gaeta, Federico  
APPLICANT: Grey, Howard M.  
APPLICANT: Sidney, John  
APPLICANT: Alexander, Jeffrey L.  
TITLE OF INVENTION: Alteration of Immune Response Using Pan  
TITLE OF INVENTION: DR-Binding Peptides  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,871A  
FILING DATE: 14-SEP-1994  
CLASSIFICATION: 424

### PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/121,101

FILING DATE: 14-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 14137-0062-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-305-871A-17

Query Match 69.4%; Score 25; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 1 WTLK 4

RESULT 2  
US-08-788-822A-1  
; Sequence 1, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/788,822A  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010,510  
; FILING DATE: 24-JAN-1996  
; NAME: Bastian, Kevin L.  
; REGISTRATION/DOCKET NUMBER: 014137-009210US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-788-822A-1

Query Match 69.4%; Score 25; DB 4; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 1 WTLK 4

RESULT 3  
US-08-788-822A-21  
; Sequence 21, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/788,822A  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010,510  
; FILING DATE: 24-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION/DOCKET NUMBER: 014137-009210US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 12  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "Xaa = alaninamide"  
US-08-788-822A-21

Query Match 69.4%; Score 25; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 7 WTLK 10

RESULT 4  
US-08-305-871A-22  
; Sequence 22, Application US/08305871A  
; Patent No 5736142  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Gaeta, Federico  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sidney, John  
; APPLICANT: Alexander, Jeffrey L.  
; TITLE OF INVENTION: Alteration of Immune Response Using Pan  
; TITLE OF INVENTION: DR-Binding Peptides  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA

```
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/305.871A
/ FILING DATE: 14-SEP-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/121,101
/ FILING DATE: 14-SEP-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 14137-0062-10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0300
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 13 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 1..13
/ OTHER INFORMATION: /note= "Peptide wherein X is
/ OTHER INFORMATION: tyrosine or phenylalanine."
US-08-305-871A-22
```

```
Query Match 69.4%; Score 25; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 7 WTLK 10
Db 7 WTLK 10
```

```
RESULT 5
US-08-305-871A-23
/ Sequence 23, Application US/08305871A
/ Patent No. 5736142
/ GENERAL INFORMATION:
/ APPLICANT: Sette, Alessandro
/ APPLICANT: Gaeta, Federico
/ APPLICANT: Grey, Howard M.
/ APPLICANT: Sidney, John
/ APPLICANT: Alexander, Jeffrey L.
/ TITLE OF INVENTION: Alteration of Immune Response Using Pan
/ TITLE OF INVENTION: DR-Binding Peptides
/ NUMBER OF SEQUENCES: 29
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/305.871A
/ FILING DATE: 14-SEP-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US 08/121,101
/ FILING DATE: 14-SEP-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 14137-0062-10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0300
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 13 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 1..13
/ OTHER INFORMATION: /note= "Peptide wherein X is
/ OTHER INFORMATION: tyrosine or phenylalanine."
US-08-305-871A-23
```

```
Query Match 69.4%; Score 25; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 7 WTLK 10
Db 7 WTLK 10
```

```
RESULT 6
US-08-788-822A-26
/ Sequence 26, Application US/08788822A
/ Patent No. 6413935
/ GENERAL INFORMATION:
/ APPLICANT: Alexander, Jeffrey L.
/ APPLICANT: DeFrees, Shawn
/ APPLICANT: Sette, Alessandro
/ TITLE OF INVENTION: Induction of Immune Response Against
/ TITLE OF INVENTION: Desired Determinants
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/788.822A
/ FILING DATE: 23-JAN-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/010,510
/ FILING DATE: 24-JAN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 014137-009210US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 13 amino acids
/ TYPE: amino acid
```

;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 3  
;; OTHER INFORMATION: /product= "OTHER"  
;; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"  
US-08-788-822A-26

Query Match 69.4%; Score 25; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||  
Db 7 WTLK 10

RESULT 7  
US-08-788-822A-27  
; Sequence 27, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/788,822A  
FILING DATE: 23-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/010,510  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 014137-009210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-788-822A-27

Query Match 69.4%; Score 25; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||  
Db 7 WTLK 10

RESULT 8  
US-08-788-822A-30  
; Sequence 30, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/788,822A  
FILING DATE: 23-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/010,510  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 014137-009210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 3  
OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /note= "Xaa = cyclohexylalanine, Tyr,  
OTHER INFORMATION: Phe or conservative substitutions  
OTHER INFORMATION: therefor"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 7..8  
OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /note= "Xaa may be present or absent"  
US-08-788-822A-30

Query Match 69.4%; Score 25; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||  
Db 9 WTLK 12

RESULT 9  
US-08-652-369A-1  
; Sequence 1, Application US/08652369A  
; Patent No. 5861268  
; GENERAL INFORMATION:

; APPLICANT: Dean G. Tang, Kenneth V. Horn  
 ; TITLE OF INVENTION: Induction of Tumor  
 ; TITLE OF INVENTION: Cell Apoptosis With Chemical Inhibitors  
 ; TITLE OF INVENTION: Targeted to 12-Lipoxygenase  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ian C. McLeod  
 ; STREET: 2190 Commons Parkway  
 ; CITY: Okemos  
 ; STATE: Michigan  
 ; COUNTRY: USA  
 ; ZIP: 48864

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette 5.25 inch,  
 ; MEDIUM TYPE: 360 kb storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS (version  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,369A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ian C. McLeod  
 ; REGISTRATION NUMBER: 20,931  
 ; REFERENCE/DOCKET NUMBER: Blomide 4.1-11  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (517) 347-4100  
 ; TELEFAX: (517) 347-4103  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 18  
 ; TYPE: Amino Acid  
 ; STRANDEDNESS: Single  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE:

; DESCRIPTION: DNA primer for segment  
 ; DESCRIPTION: of 12-lipoxygenase.  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: human  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: Genomic

; US-08-652-369A-1  
 Query Match 69.4%; Score 25; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 1 WTLK 4

; RESULT 10  
 ; US-07-741-453A-25  
 ; Sequence 25, Application US/07741453A  
 ; Patent No. 6228597

; GENERAL INFORMATION:

; APPLICANT: PARMENTIER, MARC  
 ; APPLICANT: LIBERT, FREDERIC  
 ; APPLICANT: DUMONT, JACQUES  
 ; APPLICANT: VASSART, GILBERT

; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR  
 ; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS  
 ; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES  
 ; NUMBER OF SEQUENCES: 62  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN DAREY & CUSHMAN  
 ; STREET: 1615 L STREET, N.W.  
 ; CITY: WASHINGTON, D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/741,453A  
 ; FILING DATE: 19911015  
 ; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KOKULIS, PAUL N.  
 ; REGISTRATION NUMBER: 16773  
 ; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 861-3000  
 ; TELEFAX: (202) 822-0944  
 ; TELEX: 6714627 CUSH  
 ; INFORMATION FOR SEQ ID NO: 25:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide

; US-07-741-453A-25  
 Query Match 69.4%; Score 25; DB 4; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 4 WTLK 7

; RESULT 11  
 ; US-09-075-257A-17  
 ; Sequence 17, Application US/09075257A  
 ; Patent No. 6074645

; GENERAL INFORMATION:

; APPLICANT: DIAMOND, DON JEFFREY

; APPLICANT: YORK, JOANNE

; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES

; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BART G. NEWLAND

; STREET: 555 13TH STREET, NW SUITE 701E

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/075,257A

; FILING DATE: 11-MAY-1998

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/021,298

; FILING DATE: 10-FEB-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/950,064

; FILING DATE: 14-OCT-1997

; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 08/747,488
; FILING DATE: 12-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NEWLAND, BART G
; REGISTRATION NUMBER: 31,282
; REFERENCE/DOCKET NUMBER: 1954-112CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 6
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine
; OTHER INFORMATION: or phenylalanine"
;
US-09-075-257A-17
;
Query Match 69.4%; Score 25; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 10 WTLK 13

RESULT 12
US-09-534-639-17
; Sequence 17, Application US/09534639
; Patent No. 6251399
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN
; FILE REFERENCE: 1954-343
; CURRENT APPLICATION NUMBER: US/09/534,639
; EARLIER FILING DATE: 2000-03-27
; EARLIER FILING DATE: 1998-05-11
; EARLIER APPLICATION NUMBER: 09/075,257
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: 08/950,064
; EARLIER FILING DATE: 1997-10-14
; EARLIER APPLICATION NUMBER: 08/747,488
; EARLIER FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (6)
; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human
; OTHER INFORMATION: Cytomegalovirus Vaccine
;
US-09-534-639-17
;
Query Match 69.4%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK-10
   ||||
Db 10 WTLK 13

RESULT 13
US-09-075-257A-15
; Sequence 15, Application US/09075257A
; Patent No. 6074645
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, DON JEFFREY
; APPLICANT: YORK, JOANNE
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BART G. NEWLAND
; STREET: 555 13TH STREET, NW SUITE 701E
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,257A
; FILING DATE: 11-MAY-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/021,298
; FILING DATE: 10-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/950,064
; FILING DATE: 14-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/747,488
; FILING DATE: 12-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NEWLAND, BART G
; REGISTRATION NUMBER: 31,282
; REFERENCE/DOCKET NUMBER: 1954-112CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 6
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine
; OTHER INFORMATION: or phenylalanine"
;
US-09-075-257A-15
;
Query Match 69.4%; Score 25; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 10 WTLK 13

RESULT 14
US-09-075-257A-16
; Sequence 16, Application US/09075257A
```



; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YORK, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 6  
; OTHER INFORMATION: /note= "xaa = cyclohexylalanine  
; OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-16

Query Match 69.4%; Score 25; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|  
|  
|  
|  
Db 10 WTLK 13

RESULT 15  
US-09-534-639-15  
; Sequence 15 Application US/09534639  
; Patent No. 6251393  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J  
; APPLICANT: York, Joanne  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
; TITLE OF INVENTION: CYTOMEGALOVIRUS  
; FILE REFERENCE: 1954-343

; CURRENT APPLICATION NUMBER: US/09/534,639  
; CURRENT FILING DATE: 2000-03-27  
; EARLIER APPLICATION NUMBER: 09/075,257  
; EARLIER FILING DATE: 1998-05-11  
; EARLIER APPLICATION NUMBER: 09/021,298  
; EARLIER FILING DATE: 1998-02-10  
; EARLIER APPLICATION NUMBER: 08/950,064  
; EARLIER FILING DATE: 1997-10-14  
; EARLIER APPLICATION NUMBER: 08/747,488  
; EARLIER FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 28  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Human  
; OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-15

Query Match 69.4%; Score 25; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|  
|  
|  
|  
Db 10 WTLK 13

Search completed: January 29, 2003, 10:48:22  
Job time : 10.9091 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:07 : Search time 10.8727 Seconds  
(without alignments)  
114.943 Million cell updates/sec

Title: US-09-707-738-22  
Perfect score: 56  
Sequence: 1 AKXVAATLKAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	67.9	411	2 S61245	probable virion glycoprotein M (gm) - bovine herpesvirus 1
2	37	66.1	1036	2 B69368	hypothetical prote
3	35	62.5	204	2 A53848	ubiquitin-conjugat
4	35	62.5	942	2 G83038	probable sensor/re
5	35	62.5	1492	2 T14652	protein J - Yersin
6	35	62.5	1545	2 T14966	phage lambda-reiac
7	35	62.5	2429	1 SUHUA	spectrin alpha cha
8	34	60.7	85	1 IHKREV	high potential iro
9	34	60.7	102	2 H87678	hypothetical prote
10	34	60.7	201	2 C97558	hypothetical prote
11	34	60.7	201	2 A22778	conserved hypotet
12	34	60.7	201	2 A22778	conserved hypotet
13	34	60.7	297	2 E95379	probable arginase
14	34	60.7	363	2 B82532	GTP-binding protei
15	34	60.7	418	2 C84360	potassium channel
16	34	60.7	489	2 C70655	probable monooxyge
17	34	60.7	494	2 A86917	probable monooxyge
18	34	60.7	494	2 G87567	monooxygenase, fla
19	34	60.7	499	2 B83493	conserved hypotet
20	34	60.7	537	2 S73612	arginine-tRNA liga
21	34	60.7	731	2 T09172	probable calcium-a
22	34	60.7	845	2 T34064	hypothetical prote
23	34	60.7	940	2 AB1744	internalin protein
24	33	58.9	163	2 S69658	hypothetical prote
25	33	58.9	293	2 A12626	transcription regu
26	33	58.9	295	2 B86120	3-phosphoserine ph
27	33	58.9	295	2 T51162	phosphoserine phos
28	33	58.9	299	2 H97408	probable transcrip
29	33	58.9	339	2 AE3502	D-xylose 1-dehydro

30	33	58.9	378	2 T06744	hypothetical prote
31	33	58.9	940	2 AD1374	internalin protein
32	33	58.9	970	2 D59435	Gen-interacting pr
33	33	58.9	1253	2 T21085	hypothetical prote
34	32.5	58.0	396	2 B75290	hypothetical prote
35	32	57.1	19	2 I53673	amyloid protein -
36	32	57.1	146	2 D70766	hypothetical prote
37	32	57.1	194	2 AF3408	CDPdiacylglycerol-
38	32	57.1	228	2 AG2451	4-diphosphocytidyl
39	32	57.1	233	2 A87601	hypothetical prote
40	32	57.1	246	2 AG2988	competence protein
41	32	57.1	291	2 A92955	conserved hypotet
42	32	57.1	317	2 F87336	lipoprotein (impor
43	32	57.1	330	2 G89771	para-aminobenzoate
44	32	57.1	344	2 AF2782	probable GTP-bind
45	32	57.1	367	2 F64202	

## ALIGNMENTS

## RESULT 1

S61245  
probable virion glycoprotein M (gm) - bovine herpesvirus 1  
C:Species: bovine herpesvirus 1  
C>Date: 18-Sep-1997 #sequence\_revision 18-Sep-1997 #text\_change 26-Aug-1999  
R:Accession: S61245  
R:Vicek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Schw  
submitted to the EMBL Data Library, January 1995  
A:Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1  
A:Reference number: S61233  
A:Accession: S61245  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-411 <VLC>  
A:Cross-references: EMBL:Z48053; NID:G971311; PIDN:CAA88123.1; PID:G971324  
C:Superfamily: herpesvirus 51K protein

Query Match 67.9%; Score 38; DB 2; Length 411;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12  
:|||||:  
Db 141 LAATLQAA 149

## RESULT 2

B69368  
hypothetical protein AF0946 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: B69368  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1036 <KLE>  
A:Cross-references: GB:AE001039; GB:AE000782; NID:G2689362; PIDN:AA890312.1; PID:G264967

Query Match 66.1%; Score 37; DB 2; Length 1036;  
Best Local Similarity 58.3%; Pred. No. 71;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKXVAATLKAA 12  
| | | | |

```

Db 332 ASTVGAWTVKEA 343

RESULT 3
A:Species: Pichia pastoris
A:Title: Ubiquitin-conjugating enzyme - yeast (Pichia pastoris)
A:Reference number: A53848
A:Accession: A53848
A:Reference number: A53848; MUID:94342381; PMID:8063827
A:Accession: A53848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <CRA>
A:Cross-references: GB:U12511; NID:G531386; PIDN:AAAS3634.1; PID:G531387
C:Genetics:
A:Gene: PAS4
C:Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match 62.5%; Score 35; DB 2; Length 204;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
||| ||| :|
Db 139 AKWTPAWTLSSA 150

RESULT 4
G83038
probable sensor/response regulator hybrid PA4856 [imported] - Pseudomonas aeruginosa (sp
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
A:Accession: G83038
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-942 <STO>
A:Cross-references: GB:AE004899; GB:AE004091; NID:G9951126; PIDN:AAG08241.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4856

Query Match 62.5%; Score 35; DB 2; Length 942;
Best Local Similarity 72.7%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 11
||| ||| :|
Db 765 AKPVAGVTLKA 775

RESULT 5
T14652
protein J - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
A:Accession: T14652
R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrano,
submitted to the EMBL Data Library, March 1998
A:Description: Structural organization of virulence determinants in three Yersinia pesti
A:Reference number: Z18168
A:Accession: T14652
A:Status: preliminary; translated from GB/EMBL/DBJ

Db 332 ASTVGAWTVKEA 343

A:Molecule type: DNA
A:Residues: 1-1492 <HUP>
A:Cross-references: EMBL:AF053947; NID:G2996286; PID:G2996342; PIDN:AAC13222.1
C:Genetics:
A:Genome: plasmid pMT1

Query Match 62.5%; Score 35; DB 2; Length 1492;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAAWTLK 10
||| ||| :|
Db 1248 KVSAAWTLK 1256

RESULT 6
T14966
phage lambda-related host specificity protein J - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 22-Oct-1999
A:Accession: T14966
R:Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plasmid
A:Reference number: Z18268; MUID:99043898; PMID:9826348
A:Accession: T14966
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1545 <LIN>
A:Cross-references: EMBL:AF074611; NID:G3883003; PID:G3883049; PIDN:AAC82709.1
C:Genetics:
A:Gene: Y1049
A:Genome: plasmid pMT1

Query Match 62.5%; Score 35; DB 2; Length 1545;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAAWTLK 10
||| ||| :|
Db 1289 KVSAAWTLK 1297

RESULT 7
SUHUA
spectrin alpha chain - human
C:Species: Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
A:Accession: A35716; A45755; A23533; A91528; A42872; A02965; S13138; A05282; A23965; A270
R:Sahr, K.E.; Laurila, P.; Kotula, L.; Scarpa, A.L.; Coupal, E.; Leto, T.L.; Linnenbach,
J. Biol. Chem. 265, 4434-4443, 1990
A:Title: The complete cDNA and polypeptide sequences of human erythroid alpha-spectrin.
A:Reference number: A35716; MUID:90170949; PMID:1689726
A:Accession: A35716
A:Molecule type: mRNA
A:Residues: 1-2429 <SAH>
A:Cross-references: GB:J05244
R:Sahr, K.E.; Tobe, T.; Scarpa, A.; Laughinghouse, K.; Marchesi, S.L.; Agre, P.; Linnenba
J. Clin. Invest. 84, 1243-1252, 1989
A:Title: Sequence and exon-intron organization of the DNA encoding the alphaI domain of h
A:Reference number: A45755; MUID:90009318; PMID:2794061
A:Accession: A45755
A:Molecule type: DNA
A:Residues: 1-394, 'G', 396-533 <SA2>
A:Cross-references: GB:M29983
A:Note: The authors translated the codon GGT for residue 395 as Ala
R:Linnenbach, A.J.; Speicher, D.W.; Marchesi, V.T.; Forget, B.G.
Proc. Natl. Acad. Sci. U.S.A. 83, 2397-2401, 1986
A:Title: Cloning of a portion of the chromosomal gene for human erythrocyte alpha-spectr
A:Reference number: A23533; MUID:86205962; PMID:3458204
A:Accession: A23533
A:Molecule type: DNA
A:Residues: 320-450 <LIN>

```

A;Cross-references: GB:M1233; NID:gl82242; PIDN:AA53103.1; PID:gl82243  
 R;Curtis, P.J.; Palumbo, A.; Ming, J.; Fraser, P.; Cloe, L.; Meo, P.; Shane, S.; Rovera, Gene 36, 357-362, 1985  
 A;Title: Sequence comparison of human and murine erythrocyte alpha-spectrin cDNA.  
 A;Reference number: A91528; MUID:86083178; PMID:3000887  
 A;Accession: A91528  
 A;Molecule type: mRNA  
 A;Residues: 1451-1687 <CUR>  
 A;Cross-references: GB:M1049; NID:g338310; PIDN:AAA60569.1; PID:g553648  
 R;Speicher, D.W.; Meglarz, L.; Desilva, T.M.  
 J. Biol. Chem. 267, 14775-14782, 1992  
 A;Title: Properties of human red cell spectrin heterodimer (side-to-side) assembly and J. Biol. Chem. 258, 14938-14947, 1983  
 A;Reference number: A42872; MUID:92340516; PMID:1634521  
 A;Accession: A42872  
 A;Molecule type: protein  
 A;Residues: 7-16;46-55;680-689;1047-1056;1921-1930 <SP1>  
 R;Speicher, D.W.; Davis, G.; Marchesi, V.T.  
 J. Biol. Chem. 258, 14938-14947, 1983  
 A;Title: Structure of human erythrocyte spectrin. II. The sequence of the alpha-I domain A;Reference number: A92408; MUID:84087888; PMID:6654896  
 A;Accession: A02965  
 A;Molecule type: protein  
 A;Residues: 7-601 <SPC>  
 R;Speicher, D.W.; Davis, G.; Yurchenco, P.D.; Marchesi, V.T.  
 J. Biol. Chem. 258, 14931-14937, 1983  
 A;Title: Structure of human erythrocyte spectrin. I. Isolation of the alpha-I domain and A;Reference number: S13138; MUID:84087887; PMID:6654895  
 A;Accession: S13138  
 A;Molecule type: protein  
 A;Residues: 7-92; 'X', 94-96, 110-151; 317-342; 345-366; 'Z', 368-370; 'X', 372-373; 387-434; 452-4 R;Lusitani, D.M.; Oraishat, N.; LaBrake, C.C.; Yu, R.N.; Davis, J.; Kelley, M.R.; Fung, J. Biol. Chem. 269, 25955-25958, 1994  
 A;Title: The first human spectrin structural domain begins with serine.  
 A;Reference number: A38928; MUID:95014412; PMID:7929303  
 A;Contents: annotation; Determination of structural domain  
 C;Comment: Spectrin associates with band 4.1 and actin to form the cytoskeletal superstr s of approximately 106 residues each.  
 C;Genetics:  
 A;Gene: GDB:SPTA1  
 A;Cross-references: GDB:l19601; OMIM:182860  
 A;Map position: lq21-lq21  
 C;Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectrin/ F;52-156/Domain: spectrin/dystrophin repeat homology <SP1>  
 F;157-262/Domain: spectrin/dystrophin repeat homology <SP2>  
 F;263-368/Domain: spectrin/dystrophin repeat homology <SP3>  
 F;369-474/Domain: spectrin/dystrophin repeat homology <SP4>  
 F;475-580/Domain: spectrin/dystrophin repeat homology <SP5>  
 F;581-685/Domain: spectrin/dystrophin repeat homology <SP6>  
 F;686-791/Domain: spectrin/dystrophin repeat homology <SP7>  
 F;792-897/Domain: spectrin/dystrophin repeat homology <SP8>  
 F;898-983/Domain: spectrin/dystrophin repeat homology <SP9>  
 F;984-1031/Domain: SH3 homology <SH3>  
 F;1081-1181/Domain: spectrin/dystrophin repeat homology <SP10>  
 F;1182-1287/Domain: spectrin/dystrophin repeat homology <SP11>  
 F;1288-1393/Domain: spectrin/dystrophin repeat homology <SP12>  
 F;1394-1498/Domain: spectrin/dystrophin repeat homology <SP13>  
 F;1499-1604/Domain: spectrin/dystrophin repeat homology <SP14>  
 F;1605-1710/Domain: spectrin/dystrophin repeat homology <SP15>  
 F;1711-1816/Domain: spectrin/dystrophin repeat homology <SP16>  
 F;1817-1925/Domain: spectrin/dystrophin repeat homology <SP17>  
 F;1926-2032/Domain: spectrin/dystrophin repeat homology <SP18>  
 F;2041-2146/Domain: spectrin/dystrophin repeat homology <SP19>  
 F;2155-2257/Domain: spectrin/dystrophin repeat homology <SP20>  
 F;2270-2302/Domain: calmodulin repeat homology <EF1>  
 F;2313-2345/Domain: calmodulin repeat homology <EF2>

Query Match 62.5%; Score 35; DB 1; Length 2429;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAATLK 10  
 || : ||| |

Db 1913 AKATAAWKLQ 1922  
 RESULT 8  
 IHKREV  
 A;Title: high potential iron-sulfur protein [validated] - Chromatium vinosum  
 N;Alternate names: HiPIP  
 C;Species: Chromatium vinosum  
 C;Date: 24-Apr-1984 #sequence revision 24-Oct-1997 #text\_change 15-Sep-2000  
 C;Accession: A92330; A92143; A00263  
 R;Tedo, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.  
 J. Biol. Chem. 256, 731-735, 1981  
 A;Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the puri A;Reference number: A92330; MUID:81094036; PMID:7451471  
 A;Accession: A92330  
 A;Molecule type: protein  
 A;Residues: 1-73; 'D', 75-85 <TED>  
 R;Dus, K.; Tedo, S.; Bartsch, R.G.  
 J. Biol. Chem. 248, 7318-7331, 1973  
 A;Title: The complete amino acid sequence of Chromatium high potential iron sulfur protei A;Reference number: A92143; MUID:74012043; PMID:4745771  
 A;Accession: A92143  
 A;Molecule type: protein  
 A;Residues: 1-10; 'N', 12-44; 'D', 46-85 <DUS>  
 A;Experimental source: strain D  
 R;Banci, L.; Bertini, I.; Dikly, A.; Kastrau, D.H.W.; Luchinat, C.; Sompornpisut, P. submitted to the Brookhaven Protein Data Bank, January 1995  
 A;Reference number: A65814; PDB:1HRQ  
 A;Contents: annotation; conformation by (1)H-NMR, reduced form, residues 1-85 submitted to the Brookhaven Protein Data Bank, December 1995  
 R;Bertini, I.; Dikly, A.; Kastrau, D.H.W.; Luchinat, C.; Sompornpisut, P.  
 A;Reference number: A66207; PDB:1NEH  
 A;Contents: annotation; conformation by (1)H-NMR, oxidized form, residues 1-85 R;Backes, G.; Mino, Y.; Loefer, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman, J. Am. Chem. Soc. 113, 2055-2064, 1991  
 A;Title: The environment of Fe4S4 clusters in ferredoxins and high-potential iron protein A;Reference number: A44688  
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms  
 A;Note: assignment of Raman spectra frequencies and hydrogen bonds around the iron-sulfur R;Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Xuong, N.H.; Alden, R.A.; Bartsch, R.G. J. Biol. Chem. 249, 4212-4225, 1974  
 A;Title: Two-angstrom crystal structure of oxidized Chromatium high potential iron protei A;Reference number: A92153; MUID:74309824; PMID:4855287  
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms  
 R;Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Alden, R.A.  
 J. Biol. Chem. 249, 6339-6346, 1974  
 A;Title: Comparison of oxidation-reduction site geometries in oxidized and reduced Chrome A;Reference number: A92161; MUID:75019502; PMID:4417854  
 A;Contents: annotation; X-ray crystallography  
 A;Note: structures of the oxidized and reduced forms are compared with each other and wit C;Superfamily: high potential iron-sulfur protein  
 C;Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein  
 F;43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental  
 Query Match 60.7%; Score 34; DB 1; Length 85;  
 Best Local Similarity 85.7%; Pred. No. 22;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAATLKA 11  
 : |||||  
 Db 78 ASWTLKA 84

RESULT 9  
 IHKFER  
 A;Title: high potential iron-sulfur protein - Thiocapsa roseopersicina (tentative sequence)  
 C;Species: Thiocapsa roseopersicina  
 C;Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 31-Mar-2000  
 C;Accession: A00264  
 R;Tedo, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.  
 J. Biol. Chem. 256, 731-735, 1981  
 A;Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the puri A;Reference number: A92330; MUID:81094036; PMID:7451471

A:Accession: A00264  
 A:Molecule type: protein  
 A:Residues: 1-85 <TED>  
 C:Superfamily: high potential iron-sulfur protein  
 C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein  
 F:43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 60.7%; Score 34; DB 1; Length 85;  
 Best Local Similarity 85.7%; Pred. No. 22;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAATLKA 11  
 | : |||||  
 Db 78 ASATLKA 84  
 | : |||||

RESULT 10  
 H87678  
 hypothetical protein CC3466 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: H87678  
 R:NIrman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolof  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: H87678  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-102 <STO>  
 A:Cross-references: GB:AE005673; NID:gl3425186; PIDN:AAK25428.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC3466

Query Match 60.7%; Score 34; DB 2; Length 102;  
 Best Local Similarity 58.3%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKXVAATLKAA 12  
 | : |||||  
 Db 56 ARAGAAATTKQA 67  
 | : |||||

RESULT 11  
 C97558  
 hypothetical protein AGR\_C\_3031 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C:Accession: C97558  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; PMID:11743194  
 A:Accession: C97558  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-201 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK87420.1; PID:gl5156733; GSPDB:GN00169  
 C:Genetics:  
 A:Gene: AGR\_C\_3031  
 A:Map position: circular chromosome

Query Match 60.7%; Score 34; DB 2; Length 201;  
 Best Local Similarity 50.0%; Pred. No. 51;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAATLKAA 11  
 | : |||||  
 Db 191 REIVAWTLKS 200  
 | : |||||

RESULT 12  
 AG2778  
 conserved hypothetical protein Atul644 [imported] - Agrobacterium tumefaciens (strain C58)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AG2778  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;  
 exage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krepan, W.; Perry, M.; Gordon-Kamm, E.  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AG2778  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-201 <KUR>  
 A:Cross-references: GB:AE008688; PIDN:AAI42645.1; PID:gl7740076; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atul644  
 A:Map position: circular chromosome

Query Match 60.7%; Score 34; DB 2; Length 201;  
 Best Local Similarity 50.0%; Pred. No. 51;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAATLKAA 11  
 | : |||||  
 Db 191 REIVAWTLKS 200  
 | : |||||

RESULT 13  
 E95379  
 probable arginase (EC 3.5.3.1) ArgI2 [imported] - Sinorhizobium meliloti (strain 1021) me  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 14-Sep-2001  
 C:Accession: E95379  
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse  
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
 A:Reference number: A95262; MUID:21396509; PMID:11481432  
 A:Accession: E95379  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-297 <KUR>  
 A:Cross-references: GB:AE006469; PIDN:AAK65599.1; PID:gl4524081; GSPDB:GN00165  
 A:Experimental source: strain 1021, megaplasmid pSymA  
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: argI2  
 A:Genome: plasmid  
 C:Keywords: hydrolase

Query Match 60.7%; Score 34; DB 2; Length 297;  
 Best Local Similarity 50.0%; Pred. No. 75;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKXVAATLKAA 12  
 | : |||||  
 Db 64 ARAIAAWSPKLA 75  
 | : |||||

Search completed: January 29, 2003, 10:47:34  
Job time : 11.8727 secs

## RESULT 14

B82532  
GTP-binding protein XF2641 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: B82532  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82532  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-363 <SIM>  
A:Cross-references: GB:AE004071; GB:AE003849; NID:g9107866; PIDN:AAF85438.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Canargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carrex, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2641  
C:Superfamily: yeast probable purine nucleotide-binding protein YBR025C

Query Match 60.7%; Score 34; DB 2; Length 363;  
Best Local Similarity 60.0%; Pred. No. 91;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAAWTLKA 11  
| | | | |  
DB 288 KEVRAWTVK 297

## RESULT 15

C84360  
potassium channel homolog [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: C84360  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: C84360  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-418 <STO>  
A:Cross-references: GB:AE004437; NID:g10581520; PIDN:AAG20247.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: pchB

Query Match 60.7%; Score 34; DB 2; Length 418;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 AWTLKAA 12  
| | | | |  
DB 265 AWTLRAA 271





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:30:17 ; Search time 5.90909 Seconds  
(without alignments)  
91.248 Million cell updates/sec

Title: US-09-707-738-22

Perfect score: 56  
Sequence: 1 AKXVAAMTLKAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	67.9	411	1 VGLM_HSVBC	P52370 bovine herp
2	35	62.5	204	1 UBCK_PICPA	P49428 pichia past
3	35	62.5	2418	1 SFCA_HUMAN	P02549 homo sapien
4	34	60.7	83	1 HPIS_THETI	P80176 thermochrom
5	34	60.7	85	1 HPIS_THIRO	P00261 thiocapsa r
6	34	60.7	122	1 HPIS_CHRVI	P00260 chromatiu
7	34	60.7	537	1 SYR_MYCPN	P75222 mycoplasma
8	34	60.7	574	1 KNC2_MOUSE	P58390 mus musculu
9	34	60.7	579	1 KNC2_HUMAN	Q9h2e1 homo sapien
10	34	60.7	580	1 KNC2_RAT	P70604 rattus norv
11	32	57.1	146	1 YK81_MYCTU	Q10689 mycobacteri
12	32	57.1	228	1 ISPD_ANASP	Q9y1x9 anabaena sp
13	32	57.1	367	1 Y024_MYCGE	P47270 mycoplasma
14	32	57.1	406	1 KPBH_MOUSE	Q9db30 mus musculu
15	32	57.1	406	1 KPBH_RAT	P31325 rattus norv
16	32	57.1	428	1 MTT8_THETH	P29749 thermus the
17	32	57.1	516	1 TQPL_BPT4	P09176 bacterioph
18	32	57.1	570	1 IF37_CABEL	P30642 caenorhabdi
19	32	57.1	757	1 YDK_BACSU	Q06993 bacillus su
20	32	57.1	768	1 LEM3_MOUSE	Q01102 mus musculu
21	32	57.1	770	1 A4_MOUSE	P12023 mus musculu
22	32	57.1	770	1 A4_RAT	P08592 rattus norv
23	32	57.1	776	1 TFR1_CHICK	Q90997 gallus gall
24	32	57.1	819	1 CLPC_GUITH	O78410 guillardia
25	32	57.1	821	1 CLPC_PORPU	P61332 porphyra pu
26	32	57.1	830	1 LEM3_HUMAN	P6109 homo sapien
27	32	57.1	907	1 NUOG_ECOLI	P33602 escherichia
28	32	57.1	907	1 NUOG_SALTY	P33900 salmonella
29	32	57.1	2241	1 TEGU_HCMVA	P16785 human cytom
30	31	55.4	83	1 HPIS_CHRGR	P00262 chromatiu
31	31	55.4	250	1 TRMD_TREPA	O83878 treponema p
32	31	55.4	260	1 CAID_ECO57	Q8x435 escherichia
33	31	55.4	260	1 CAID_ECOLI	P31551 escherichia

34	31	55.4	260	1 CAID_SALTY	Q8z9l5 salmonella
35	31	55.4	260	1 CAID_SALTY	Q8zrx5 salmonella
36	31	55.4	320	1 MTD1_YEAST	Q02046 saccharomyc
37	31	55.4	363	1 ALFB_HUMAN	P05062 homo sapien
38	31	55.4	363	1 ALFB_RAT	P00884 rattus norv
39	31	55.4	378	1 PLYB_ASPNG	Q00205 aspergillus
40	31	55.4	379	1 SGAA_METEX	P55819 methylobact
41	31	55.4	406	1 KPBH_HUMAN	P15735 homo sapien
42	31	55.4	472	1 BMP3_HUMAN	P12645 homo sapien
43	31	55.4	486	1 ALLN_ALISA	Q01594 allium sati
44	31	55.4	521	1 FD33_MYCTU	Q11015 mycobacteri
45	31	55.4	536	1 KNC1_RAT	P70606 rattus norv

ALIGNMENTS

```
RESULT 1
VGLM_HSVBC
ID VGLM_HSVBC STANDARD; PRT; 411 AA.
AC P52370;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein M.
GN GM OR UL10.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=953113343; PubMed=7793062;
RA Vleck C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyzler M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 248053; CAA88123.1; -.
CC InterPro; IPR000785; Herpes_glycop.
CC Pfam; PF01528; Herpes_glycop; 1.
CC PRINTS; PR00333; HSVINTEGRIMP.
CC Transmembrane; Glycoprotein.
CC TRANSMEM 14 34 POTENTIAL.
CC TRANSMEM 89 109 POTENTIAL.
CC TRANSMEM 133 153 POTENTIAL.
CC TRANSMEM 159 179 POTENTIAL.
CC TRANSMEM 213 233 POTENTIAL.
CC TRANSMEM 249 269 POTENTIAL.
CC TRANSMEM 277 297 POTENTIAL.
CC TRANSMEM 318 338 POTENTIAL.
CC CARBOHYD 57 57 N-LINKED (GLCNAC...) (POTENTIAL).
CC SEQUENCE 411 AA; 43029 MW; 20F156DA9F40158C CRC64;
Query Match 67.9%; Score 38; DB 1; Length 411;
Best Local Similarity 77.8%; Pred. No. 7.3;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

Qy 4 VAAWTLKAA 12  
Db 141 LAAWTLQAA 149

```

RESULT 2
UBCX_PICPA
ID  UBCX_PICPA  STANDARD;  PRT;  204 AA.
AC  P49428;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Ubiquitin-conjugating enzyme E2-24 kDa (EC 6.3.2.19) (Ubiquitin-
DE  protein ligase) (Ubiquitin carrier protein) (Peroxin-4).
GN  PE44 OR PAS4.
OS  Pichia pastoris (Yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Pichia.
OX  NCBI_TaxID=4922;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94342381; PubMed=8063827;
RA  Crane D.R., Kalish J.E., Gould S.J.;
RT  "The Pichia pastoris PAS4 gene encodes a ubiquitin-conjugating enzyme
RT  required for peroxisome assembly."
RL  J. Biol. Chem. 269:21835-21844(1994).
CC  -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC  OTHER PROTEINS. ESSENTIAL FOR PEROXISOME BIOGENESIS.
CC  -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC  diphosphate + protein N-ubiquityllysine.
CC  -!- PATHWAY: Ubiquitin conjugation; second step.
CC  -!- SUBCELLULAR LOCATION: Peroxisomal.
CC  -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC  UBIQUITIN-THIOLESTER FORMATION.
CC  -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC  STRONGEST, TO YEAST UBC10.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sb-sib.ch)
CC  -----
DR  EMBL; U12511; AA553634.1; -.
DR  HSP; Q95044; 2EXC.
DR  InterPro; IPR000608; UBQ_conjugat.
DR  Pfam; PF00179; UQ_con; 1.
DR  ProDom; PD000461; UBQ_conjugat; 1.
DR  SMART; SM00212; UBQC; 1.
DR  PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR  PROSITE; PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW  Ub1 conjugation pathway; Ligase; Multigene family; Peroxisome.
FT  BINDING 133 133 UBIQUITIN (BY SIMILARITY).
FT  MUTAGEN 133 133 C-S, A: LOSS OF ACTIVITY.
SQ  SEQUENCE 204 AA; 23565 MW; 46680ABD73121F6C CRC64;

Query Match 62.5%; Score 35; DB 1; Length 204;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy  1 AKXVAATLKAA 12
Db  139 AKWTPAWTLSSA 150

RESULT 3
SPCA HUMAN
ID  SPCA HUMAN  STANDARD;  PRT;  2418 AA.
AC  P02549; Q15514;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Spectrin alpha chain, erythrocyte (Erythroid alpha-spectrin).
GN  SPTA1 OR SPTA.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90170949; PubMed=1689726;
RA  Sahr K.E., Laurila P., Kotula L., Scarpa A.L., Coupal E., Leto T.L.,
RA  Linnenbach A.J., Winkelmann J.C., Speicher D.W., Marchesi V.T.,
RA  Curtis P.J., Forget B.G.;
RT  "The complete cDNA and polypeptide sequences of human erythroid
RT  alpha-spectrin."
RL  J. Biol. Chem. 265:4434-4443(1990).
RN  [2]
RP  SEQUENCE OF 7-533 FROM N.A., AND VARIANTS PRO-260; PRO-261 AND
RP  PRO-471.
RX  MEDLINE=90009318; PubMed=2794061;
RA  Sahr K.E., Tobe T., Scarpa A., Laughinghouse K., Marchesi S.L.,
RA  Agre P., Linnenbach A.J., Marchesi V.T., Forget B.G.;
RT  "Sequence and exon-intron organization of the DNA encoding the alpha
RT  I domain of human spectrin. Application to the study of mutations
RT  causing hereditary elliptocytosis."
RL  J. Clin. Invest. 84:1243-1252(1989).
RN  [3]
RP  SEQUENCE OF 7-601.
RX  MEDLINE=84087888; PubMed=6654896;
RA  Speicher D.W., Davis G., Marchesi V.T.;
RT  "Structure of human erythrocyte spectrin. II. The sequence of the
RT  alpha-I domain."
RL  J. Biol. Chem. 258:14938-14947(1983).
RN  [4]
RP  SEQUENCE OF 7-125.
RX  MEDLINE=84087887; PubMed=6654895;
RA  Speicher D.W., Davis G., Yurchenco P.D., Marchesi V.T.;
RT  "Structure of human erythrocyte spectrin. I. Isolation of the alpha-I
RT  domain and its cyanogen bromide peptides."
RL  J. Biol. Chem. 258:14931-14937(1983).
RN  [5]
RP  SEQUENCE OF 320-450 FROM N.A.
RX  MEDLINE=86205962; PubMed=3458204;
RA  Linnenbach A.J., Speicher D.W., Marchesi V.T., Forget B.G.;
RT  "Cloning of a portion of the chromosomal gene for human erythrocyte
RT  alpha-spectrin by using a synthetic gene fragment."
RL  Proc. Natl. Acad. Sci. U.S.A. 83:2397-2401(1986).
RN  [6]
RP  PARTIAL SEQUENCE.
RX  MEDLINE=84295638; PubMed=6472478;
RA  Speicher D.W., Marchesi V.T.;
RT  "Erythrocyte spectrin is comprised of many homologous triple helical
RT  segments."
RL  Nature 311:177-180(1984).
RN  [7]
RP  IDENTIFICATION OF PROBABLE FRAMESHIFT IN 2407-2418.
RA  Gibson T.J.;
RL  Unpublished observations (MAR-1995).
RN  [8]
RP  REVIEW ON VARIANTS
RX  MEDLINE=97001215; PubMed=8844207;
RA  Maillet P., Alloisio N., Morle L., Delaunay J.;
RT  "Spectrin mutations in hereditary elliptocytosis and hereditary
RT  spherocytosis."
RL  Hum. Mutat. 8:97-107(1996).
RN  [9]
RP  VARIANT SER-24.
RX  MEDLINE=94289716; PubMed=8018926;
RA  Parquet N., Devaux I., Boulanger L., Galand C., Boivin P.,
RA  Lecomte M.-C., Dhery D., Garbarz M.;
RT  "Identification of three novel spectrin alpha I/74 mutations in
RT  hereditary elliptocytosis: further support for a triple-stranded
RT  folding unit model of the spectrin heterodimer contact site."
RL  Blood 84:303-308(1994).
RN  [10]
RP  VARIANTS CYS-28; HIS-28; LEU-28 AND SER-28.

```

RX MEDLINE=91358728; PubMed=1679439;  
 RA Coetzer T.L., Sahr K., Prchal J., Blacklock H., Peterson L., Koler R.,  
 RA Doyle J., Manaster J., Palek J.;  
 RT "Four different mutations in codon 28 of alpha spectrin are  
 RT associated with structurally and functionally abnormal spectrin alpha  
 RT 1/74 in hereditary elliptocytosis.";  
 RL J. Clin. Invest. 88:743-749(1991).  
 RN [11]  
 RP VARIANTS SER-28 AND ARG-48.  
 RX MEDLINE=91346849; PubMed=1878597;  
 RA Floyd P.B., Gallagher P.G., Valentino L.A., Davis M., Marchesi S.L.,  
 RA Forget B.G.;  
 RT "Heterogeneity of the molecular basis of hereditary  
 RT pyropoikilocytosis and hereditary elliptocytosis associated with  
 RT increased levels of the spectrin alpha I/74-kilodalton tryptic  
 RT peptide.";  
 RL Blood 78:1364-1372(1991).  
 RN [12]  
 RP VARIANTS SER-45.  
 RX MEDLINE=89323468; PubMed=2568862;  
 RA Lecomte M.-C., Garbarz M., Grandchamp B., Feo C., Gautero H.,  
 RA Devaux I., Bournier O., Galand C., D'Auriol L., Galibert F.,  
 RA Sahr K.E., Forget B.G., Boivin P., Dhermy D.;  
 RT "Sp alpha I/78: a mutation of the alpha I spectrin domain in a white  
 RT kindred with HE and HPP phenotypes.";  
 RL Blood 74:1126-1133(1989).  
 RN [13]  
 RP VARIANTS PRO-207.  
 RX MEDLINE=92176375; PubMed=1541680;  
 RA Gallagher P.G., Tse W.T., Coetzer T., Lecomte M.-C., Garbarz M.,  
 RA Zarkowsky H.S., Baruchel A., Ballas S.K., Dhermy D., Palek J.,  
 RA Forget B.G.;  
 RT "A common type of the spectrin alpha I 46-50a-kD peptide abnormality  
 RT in hereditary elliptocytosis and pyropoikilocytosis is associated  
 RT with a mutation distant from the proteolytic cleavage site. Evidence  
 RT for the functional importance of the triple helical model of  
 RT spectrin.";  
 RL J. Clin. Invest. 89:892-898(1992).  
 RN [14]  
 RP VARIANTS VAL-1857.  
 RX MEDLINE=93253053; PubMed=8486776;  
 RA Wilmotte R., Marechal J., Morle L., Baklouti F., Philippe N.,  
 RA Kastally R., Kotula L., Delaunay J., Alloisio N.;  
 RT "Low expression allele alpha IELF of red cell spectrin is associated  
 RT with mutations in exon 40 (alpha V/41 polymorphism) and intron 45 and  
 RT with partial skipping of exon 46.";  
 RL J. Clin. Invest. 91:2091-2096(1993).  
 RN [15]  
 RP VARIANTS BARCELONA.  
 RX MEDLINE=91372367; PubMed=8364215;  
 RA dalla Venezia N., Alloisio N., Forissier A., Denoroy L., Aymerich M.,  
 RA Vives-Corrons J.L., Besalduch J., Besson I., Delaunay J.;  
 RT "Ellipsoikilocytosis associated with the alpha 469 His-->Pro  
 RT mutation in spectrin Barcelona (alpha I/50-46b).";  
 RL Blood 82:1661-1665(1993).  
 RN [16]  
 RP VARIANTS CAGLIARI.  
 RX MEDLINE=94043025; PubMed=8226774;  
 RA Sahr K.E., Coetzer T.L., Moy L.S., Derick L.H., Chishci A.H.,  
 RA Jarolin P., Lorenzo F., Miraglia del Giudice E., Iolascon A.,  
 RA Gallanello R.;  
 RT "Spectrin Cagliari: an Ala-->Gly substitution in helix 1 of beta  
 RT spectrin repeat 17 that severely disrupts the structure and self-  
 RT association of the erythrocyte spectrin heterodimer.";  
 RL J. Biol. Chem. 268:22656-22662(1993).  
 RN [17]  
 RP VARIANTS CULOZ AND LYON.  
 RX MEDLINE=90347052; PubMed=2384601;  
 RA Morle L., Roux A.-F., Alloisio N., Pothier B., Starck J., Denoroy J.,  
 RA Morle F., Rudigoz R.-C., Forget B.G., Delaunay J., Godet J.;  
 RT "Two elliptocytogenic alpha I/74 variants of the spectrin alpha I  
 RT domain. Spectrin Culoz (GGT-->GTT; alpha I 40 Gly-->Val) and  
 RT spectrin Lyon (CTT-->TTT; alpha I 43 Leu-->Phe).";

RL J. Clin. Invest. 86:548-554(1990).  
 RN [18]  
 RP VARIANTS JENDOUBA.  
 RX MEDLINE=92345619; PubMed=1638030;  
 RA Alloisio N., Wilmotte R., Morle L., Baklouti F., Marechal J.,  
 RA Ducluzeau M.-T., Denoroy L., Feo C., Forget B.G., Kastally R.,  
 RA Delaunay J.;  
 RT "Spectrin Jendoubai: an alpha II/31 spectrin variant that is  
 RT associated with elliptocytosis and carries a mutation distant from  
 RT the dimer self-association site.";  
 RL Blood 80:809-815(1992).  
 RN [19]  
 RP VARIANTS TUNIS.  
 RX MEDLINE=89323436; PubMed=2568861;  
 RA Morle L., Roux A.F., Godet J., Forget B.G., Denoroy L.,  
 RA Garbarz M., Dhermy D., Kastally R., Delaunay J.;  
 RT "Spectrin Tunis (Sp alpha I/78), an elliptocytogenic variant, is due  
 RT to the CCG-->TGG codon change (Arg-->Trp) at position 35 of the  
 RT alpha I domain.";  
 RL Blood 74:828-832(1989).  
 RN [20]  
 RP VARIANTS GENOVA.  
 RX MEDLINE=94250920; PubMed=8193371;  
 RA Perrotta S., del Giudice E.M., Alloisio N., Sciaratta G., Pinto L.,  
 RA Delaunay J., Cuttillo S., Iolascon A.;  
 RT "Wild elliptocytosis associated with the alpha 34 Arg-->Trp mutation  
 RT in spectrin Genova (alpha I/74).";  
 RL Blood 83:3346-3349(1994).  
 RN [21]  
 RP VARIANTS ANASTASIA.  
 RX MEDLINE=95290423; PubMed=7772539;  
 RA Perrotta S., Iolascon A., de Angelis F., Pagano L., Colonna G.,  
 RA Cuttillo S., del Giudice E.M.;  
 RT "Spectrin Anastasia (alpha I/78): a new spectrin variant (alpha 45  
 RT Arg-->Thr) with moderate elliptocytogenic potential.";  
 RL Br. J. Haematol. 89:933-936(1995).  
 CC -1- FUNCTION: SPECTRIN IS THE MAJOR CONSTITUENT OF THE CYTOSKELETAL  
 CC NETWORK UNDERLYING THE ERYTHROCYTE PLASMA MEMBRANE. IT ASSOCIATES  
 CC WITH BAND 4.1 AND ACTIN TO FORM THE CYTOSKELETAL SUPERSTRUCTURE OF  
 CC THE ERYTHROCYTE PLASMA MEMBRANE.  
 CC -1- SUBUNIT: COMPOSED OF NONHOMOLOGOUS CHAINS, ALPHA AND BETA, WHICH  
 CC AGGREGATE SIDE-TO-SIDE IN AN ANTIPARALLEL FASHION TO FORM DIMERS,  
 CC TETRAMERS, AND HIGHER POLYMERS.  
 CC -1- DISEASE: HEREDITARY ELLIPTOCYTOSIS (HE) IS A HETEROGENEOUS  
 CC DISORDER CHARACTERIZED BY VARIABLE HEMOLYTIC ANEMIA AND ELLIPTICAL  
 CC RED CELL SHAPE. SEVERAL ABNORMALITIES IN THE MEMBRANE SKELETON  
 CC HAVE BEEN IDENTIFIED IN HE, INCLUDING A NUMBER THAT HAVE BEEN

Query Match 62.5%; Score 35; DB 1; Length 2418;  
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKXVAATLTK 10  
 |||:|||||:  
 Db 1913 AKAIAAWKLQ 1922

## RESULT 4

ID HPIS\_THETI STANDARD; PRT; 83 AA.  
 AC P80176;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein (HiPIP).  
 GN HIP.  
 OS Thermochromatium tepidum (Chromatium tepidum).  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Thermochromatium.  
 OX NCBI\_TaxID=1050;  
 RN [1]\_TaxID=1050;  
 RP SEQUENCE.  
 RC STRAIN=ATCC 43061;

```

RX MEDLINE=93343628; PubMed=8393645;
RA Moulis J.-M., Scherrer N., Gagnon J., Forest E., Petillot Y.,
RA Garcia D.;
RT "Primary structure of Chromatium tepidum high-potential iron-sulfur
RT protein in relation to thermal denaturation.";
RL Arch. Biochem. Biophys. 305:186-192(1993).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=2058546; PubMed=11095707;
RA Nogi T., Fathir I., Kobayashi M., Nozawa T., Miki K.;
RT "Crystal structures of photosynthetic reaction center and high-
RT potential iron-sulfur protein from Thermochromatium tepidum:
RT thermostability and electron transfer.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13561-13566(2000).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (0.8 ANGSTROMS).
RX MEDLINE=22071404; PubMed=12077426;
RA Liu L., Nogi T., Kobayashi M., Nozawa T., Miki K.;
RT "Ultra-high-resolution structure of high-potential iron-sulfur
RT protein from Thermochromatium tepidum.";
RL Acta Crystallogr. D 58:1085-1091(2002).
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +323 MV.
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- MISCELLANEOUS: THIS PROTEIN IS THERMOSTABLE.
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HIPIP) FAMILY.
DR PIR: S35586; S35586.
DR PDB: 1EYT: 13-DEC-00.
DR PDB: 1IUA: 20-MAR-02.
DR InterPro: IPR000170; Hipot_ironsul.
DR Pfam: PF01355; HIPIP; 1.
DR PRINTS: PR00374; HIP1PFRD0XIN.
DR PROSITE: PS00596; HIPIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S; 3D-structure.
FT METAL 43 43 IRON-SULFUR (4FE-4S).
FT METAL 46 46 IRON-SULFUR (4FE-4S).
FT METAL 61 61 IRON-SULFUR (4FE-4S).
FT METAL 75 75 IRON-SULFUR (4FE-4S).
SQ SEQUENCE 83 AA; 8786 MW; 92116E4FD2C44E0A CRC64;

Query Match 60.7%; Score 34; DB 1; Length 83;
Best Local Similarity 85.7%; Pred. No. 8.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAWTLKA 11
|:|||||
DB 76 ASWTLKA 82

RESULT 5
HPIS_THIRO STANDARD; PRT; 85 AA.
AC P00261.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High potential iron-sulfur protein (HiPIP).
GN HIP.
OS Thiocapsa roseopersicina.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OX NCBI_TaxID=1058;
RN [1]
RP SEQUENCE.
RX MEDLINE=81094036; PubMed=7451471;
RA Tedro S.M., Meyer T.E., Bartsch R.G.; Kamen M.D.;
RT "Primary structures of high potential, four-iron-sulfur ferredoxins
RT from the purple sulfur photosynthetic bacteria, Thiocapsa
RT roseopersicina and Chromatium gracile.";
RL J. Biol. Chem. 256:731-735(1981).
RN [4]
RP SEQUENCE.
RX MEDLINE=81094036; PubMed=7451471;
RA Tedro S.M., Meyer T.E., Bartsch R.G.; Kamen M.D.;
RT "Primary structures of high potential, four-iron-sulfur ferredoxins
RT from the purple sulfur photosynthetic bacteria, Thiocapsa
RT roseopersicina and Chromatium gracile.";
RL J. Biol. Chem. 256:731-735(1981).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=74309824; PubMed=485287;
RA Carter C.W. Jr., Kraut J., Freer S.T., Xuong N.H., Alden R.A.,
RA Bartsch R.G.;
RT "2-A crystal structure of oxidized Chromatium high potential iron

```

```

CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HIPIP) FAMILY.
DR PIR: A00264; IHTFER.
DR HSP: P00260; LCKU.
DR InterPro: IPR000170; Hipot_ironsul.
DR Pfam: PF01355; HIPIP; 1.
DR PRINTS: PR00374; HIP1PFRD0XIN.
DR PROSITE: PS00596; HIPIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S.
FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 46 46 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 63 63 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 77 77 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 85 AA; 8989 MW; C47AD747D2218482 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 85;
Best Local Similarity 85.7%; Pred. No. 8.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAWTLKA 11
|:|||||
DB 78 ASWTLKA 84

RESULT 6
HPIS_CHRVI STANDARD; PRT; 122 AA.
ID P00260; P96753; Q9R4K3;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High potential iron-sulfur protein precursor (HiPIP).
GN HIP.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OX Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=97320627; PubMed=9177478;
RA Bruser T., Truper H.G., Dahl C.;
RT "Cloning and sequencing of the gene encoding the high potential iron-
RT sulfur protein (HiPIP) from the purple sulfur bacterium Chromatium
RT vinosum.";
RL Biochim. Biophys. Acta 1352:18-22(1997).
RN [2]
RP SEQUENCE OF 38-122.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=74012043; PubMed=4745771;
RA Dus K., Tedro S., Bartsch R.G.;
RT "The complete amino acid sequence of Chromatium high potential iron
RT sulfur protein.";
RL J. Biol. Chem. 248:7318-7331(1973).
RN [3]
RP REVISIONS TO 48; 82 AND 111.
RX MEDLINE=81094036; PubMed=7451471;
RA Tedro S.M., Meyer T.E., Bartsch R.G.; Kamen M.D.;
RT "Primary structures of high potential, four-iron-sulfur ferredoxins
RT from the purple sulfur photosynthetic bacteria, Thiocapsa
RT roseopersicina and Chromatium gracile.";
RL J. Biol. Chem. 256:731-735(1981).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=74309824; PubMed=485287;
RA Carter C.W. Jr., Kraut J., Freer S.T., Xuong N.H., Alden R.A.,
RA Bartsch R.G.;
RT "2-A crystal structure of oxidized Chromatium high potential iron

```

RL protein.";  
 RT J. Biol. Chem. 249:4212-4225(1974).  
 [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92135210; PubMed=1734968;  
 RA Nettlesheim D.G., Harder S.R., Feinberg B.A., Otvos J.D.;  
 RT "Sequential resonance assignments of the 1H nuclear magnetic resonance  
 iron-sulfur protein from Chromatium vinosum.";  
 RL Biochemistry 31:1234-1244(1992).  
 [6]  
 RN STRUCTURE BY NMR.  
 RP MEDLINE=92304939; PubMed=1610810;  
 RX Gaillard J., Albrand J.-P., Moulis J.-M., Wemmer D.E.;  
 RA "Sequence-specific assignments of the 1H nuclear magnetic resonance  
 spectra of reduced high-potential ferredoxin (HiPiP) from Chromatium  
 vinosum.";  
 RL Biochemistry 31:5632-5639(1992).  
 [7]  
 RN STRUCTURE BY NMR.  
 RP MEDLINE=95118989; PubMed=7819198;  
 RX Banci L., Bertini I., Dikii A., Krastra D.H.W., Luchinat C.,  
 RA Sompornpisut P.;  
 RT "The three-dimensional solution structure of the reduced high-  
 potential iron-sulfur protein from Chromatium vinosum through NMR.";  
 RL Biochemistry 34:206-219(1995).  
 [8]  
 RN STRUCTURE BY NMR.  
 RP MEDLINE=95359150; PubMed=7632685;  
 RX Bertini I., Dikii A., Krastra D.H., Luchinat C., Sompornpisut P.;  
 RA "Three-dimensional solution structure of the oxidized high potential  
 iron-sulfur protein from Chromatium vinosum through NMR. Comparative  
 analysis with the solution structure of the reduced species.";  
 RL Biochemistry 34:9851-9858(1995).  
 [9]  
 RN STRUCTURE BY NMR OF MUTANT SER-114.  
 RP MEDLINE=96216867; PubMed=8639555;  
 RX Bentrop D., Bertini I., Capozzi F., Dikii A., Eltis L., Luchinat C.;  
 RA "Three-dimensional structure of the reduced C77S mutant of the  
 Chromatium vinosum high-potential iron-sulfur protein through nuclear  
 magnetic resonance; comparison with the solution structure of the  
 wild-type protein.";  
 RL Biochemistry 35:5928-5936(1996).  
 [10]  
 RN X-RAY CRYSTALLOGRAPHY (0.93 ANGSTROMS).  
 RP MEDLINE=20003126; PubMed=10531472;  
 RX Parisini E., Capozzi F., Lubini P., Lamzin V., Luchinat C.,  
 RA Sheldrick G.M.;  
 RT "Ab initio solution and refinement of two high-potential iron protein  
 structures at atomic resolution.";  
 RL Acta Crystallogr D 55:1773-1784(1999).  
 CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
 CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
 CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +360 MV.  
 CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 CC (HIPIP) FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U81381; AAB48829.1; .  
 DR PIR; A00263; IHKREV.  
 DR PDB; 1HIP; 15-APR-93.  
 DR PDB; 1HRO; 03-JUN-95.  
 DR PDB; 1HRR; 31-JUL-95.

DR PDB; 1NEH; 08-MAR-96.  
 DR PDB; 1NOE; 10-JUN-96.  
 DR PDB; 1BOY; 16-DEC-98.  
 DR PDB; 1CKU; 13-MAY-99.  
 DR InterPro; IPR00170; Hipot\_iron\_sulf.  
 DR Pfam; PF01355; HIPIP; 1.  
 DR PRINTS; PRO0374; HIPIPROXIN.  
 DR PROSITE; PS00596; HIPIP; 1.  
 KW Electron transport; Iron-sulfur; 4Fe-4S; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 37  
 FT CHAIN 38 122 HIGH POTENTIAL IRON-SULFUR PROTEIN.  
 FT METAL 80 80 IRON-SULFUR (4FE-4S).  
 FT METAL 83 83 IRON-SULFUR (4FE-4S).  
 FT METAL 100 100 IRON-SULFUR (4FE-4S).  
 FT METAL 114 114 IRON-SULFUR (4FE-4S).  
 FT CONFLICT 111 111 N -> D (IN REF. 3).  
 FT TURN 41 42  
 FT STRAND 43 43  
 FT TURN 46 47  
 FT HELIX 49 54  
 FT TURN 55 55  
 FT STRAND 57 57  
 FT HELIX 60 62  
 FT HELIX 65 68  
 FT HELIX 75 77  
 FT HELIX 80 82  
 FT STRAND 83 83  
 FT TURN 84 85  
 FT STRAND 86 87  
 FT TURN 88 89  
 FT TURN 91 92  
 FT STRAND 97 100  
 FT TURN 101 102  
 FT TURN 104 105  
 FT STRAND 107 109  
 FT TURN 110 111  
 FT STRAND 113 113  
 FT TURN 115 116  
 FT STRAND 119 119  
 SQ SEQUENCE 122 AA; 12761 MW; ACBBAPC917F32A09 CRC64;  
 Query Match 60.7%; Score 34; DB 1; Length 122;  
 Best Local Similarity 85.7%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 AAWTLKA 11  
 Db 115 ASWTLKA 121  
 .|.|.|.|  
 RESULT 7  
 SYR MYCPN STANDARD; PRT; 537 AA.  
 ID SYR MYCPN  
 AC P7522;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ArgRS).  
 GN ARGS OR MPN556 OR MP286.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +

CC    diphosphate + L-arginyl-tRNA(Arg).

CC    -!- SUBUNIT: MONOMER (BY SIMILARITY).

CC    -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC    -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC    -----

CC    This SWISS-PROT entry is copyright. It is produced through a collaboration

CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC    the European Bioinformatics Institute. There are no restrictions on its

CC    use by non-profit institutions as long as its content is in no way

CC    modified and this statement is not removed. Usage by and for commercial

CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC    -----

DR    EMBL; AE000027; AAB95934.1; -.

DR    InterPro; IPR001278; Arg\_trna-synt\_1c.

DR    InterPro; IPR005148; N.

DR    InterPro; IPR001412; trna-synt\_1.

DR    Pfam; PF00750; trna-synt\_1g; 1.

DR    Pfam; PF03485; N-Arg; 1.

DR    PRINTS; PR01038; TRNASYNTHARG.

DR    TIGRFAMs; TIGR00456; args; 1.

DR    PROSITE; PS00178; AA TRNA LIGASE I; 1.

KW    Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW    Complete proteome.

FT    SITE    113    123    "HIGH" REGION.

FT    SITE    365    369    "KMSKS" REGION.

FT    BINDING    368    368    ATP (BY SIMILARITY).

SQ    SEQUENCE    537 AA; 62120 MW; 099B1C0DC33A39B2 CRC64;

Query Match    60.7%; Score 34; DB 1; Length 537;

Best Local Similarity    60.0%; Pred. No. 53;

Matches    6; Conservative    2; Mismatches    2; Indels    0; Gaps    0;

QY    1 AKXVAATLTK 10

DB    54 AQQAATLTK 63

RESULT 8

KCN2\_MOUSE

ID    KCN2\_MOUSE    STANDARD;    PRT;    574 AA.

AC    P58350;

DT    15-JUN-2002 (Rel. 41, Created)

DT    15-JUN-2002 (Rel. 41, Last sequence update)

DT    15-JUN-2002 (Rel. 41, Last annotation update)

DE    Small conductance calcium-activated potassium channel protein 2 (SK2).

GN    KCN2 OR SK2.

OS    Mus musculus (Mouse).

OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX    NCBI\_TaxID=10090;

RN    [1]

RP    SEQUENCE FROM N.A.

RC    STRAIN=BALB/c; TISSUE=Colon;

RX    MEDLINE=21440983; PubMed=11557517;

RA    Ro S., Hattori W.J., Koh S.D., Horowitz B.;

RT    "Molecular properties of small-conductance Ca2+-activated K+ channels

RT    expressed in murine colonic smooth muscle.";

RL    Am. J. Physiol. 281:G964-G973(2001).

CC    -!- FUNCTION: FORMS A VOLTAGE-INDEPENDENT POTASSIUM CHANNEL ACTIVATED

CC    BY INTRACELLULAR CALCIUM. ACTIVATION IS FOLLOWED BY MEMBRANE

CC    HYPERPOLARIZATION. THOUGHT TO REGULATE NEURONAL EXCITABILITY BY

CC    CONTRIBUTING TO THE SLOW COMPONENT OF SYNAPTIC

CC    AFTERHYPERPOLARIZATION. THE CHANNEL IS BLOCKED BY APAMIN (BY

CC    SIMILARITY).

CC    -!- SUBUNIT: HETERO-OLIGOMER. THE COMPLEX IS COMPOSED OF 4 CHANNEL

CC    SUBUNITS EACH OF WHICH BINDS TO A CALMODULIN SUBUNIT WHICH

CC    REGULATES THE CHANNEL ACTIVITY THROUGH CALCIUM-BINDING (BY

CC    SIMILARITY).

CC    -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC    -!- SIMILARITY: BELONGS TO THE KCN FAMILY OF POTASSIUM CHANNELS.

CC    -----

CC    This SWISS-PROT entry is copyright. It is produced through a collaboration

CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC    the European Bioinformatics Institute. There are no restrictions on its

CC    use by non-profit institutions as long as its content is in no way

CC    modified and this statement is not removed. Usage by and for commercial

CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC    -----

DR    EMBL; AF357240; AAK48901.1; -.

DR    MGD; MG1:2153182; Kcnn2.

DR    InterPro; IPR004178; CaMBD.

DR    InterPro; IPR001622; K+channel\_pore.

DR    InterPro; IPR003931; SK\_channel.

DR    Pfam; PF02888; CaMBD; 1.

DR    Pfam; PF03530; SK\_channel; 1.

KW    Ionic channel; Transmembrane; Ion transport; Calmodulin-binding.

FT    TRANSMEM    133    153    SEGMENT S1 (POTENTIAL).

FT    TRANSMEM    163    183    SEGMENT S2 (POTENTIAL).

FT    TRANSMEM    209    229    SEGMENT S3 (POTENTIAL).

FT    TRANSMEM    251    271    SEGMENT S4 (POTENTIAL).

FT    TRANSMEM    300    320    SEGMENT S5 (POTENTIAL).

FT    TRANSMEM    340    360    SEGMENT H5 (FORE-FORMING) (POTENTIAL).

FT    TRANSMEM    369    389    SEGMENT S6 (POTENTIAL).

FT    DOMAIN    407    483    CALMODULIN-BINDING (BY SIMILARITY).

FT    DOMAIN    42    46    POLY-GLY.

FT    DOMAIN    52    56    POLY-ALA.

FT    DOMAIN    80    85    POLY-GLY.

FT    DOMAIN    88    97    POLY-GLY.

FT    DOMAIN    558    561    POLY-ARG.

SQ    SEQUENCE    574 AA; 63482 MW; 259CB8CDE4D1CDE6 CRC64;

Query Match    60.7%; Score 34; DB 1; Length 574;

Best Local Similarity    62.5%; Pred. No. 56;

Matches    5; Conservative    3; Mismatches    0; Indels    0; Gaps    0;

QY    4 VAAWTLKA 11

DB    319 IAAWTVRA 326

RESULT 9

KCN2\_HUMAN

ID    KCN2\_HUMAN    STANDARD;    PRT;    579 AA.

AC    Q9H2S1.

DT    15-JUN-2002 (Rel. 41, Created)

DT    15-JUN-2002 (Rel. 41, Last sequence update)

DT    15-JUN-2002 (Rel. 41, Last annotation update)

DE    Small conductance calcium-activated potassium channel protein 2 (SK2).

GN    KCN2.

OS    Homo sapiens (Human).

OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX    NCBI\_TaxID=9606;

RN    [1]

RP    SEQUENCE FROM N.A.

RX    MEDLINE=20568244; PubMed=10991935;

RA    Desai R., Peretz A., Idelson H., Lazarovici P., Attali B.;

RT    "Ca2+-activated K+ channels in human leukemic Jurkat T cells.

RT    Molecular cloning, biochemical and functional characterization.";

RL    J. Biol. Chem. 275:39954-39963(2000).

RN    [2]

RP    SEQUENCE FROM N.A.

RC    TISSUE=Myometrium;

RA    Mazzone J.N., Kaiser R.A., Buxton I.L.O.;

RT    "Characterization of calcium-activated potassium channels in human

RT    myometrium.";

RL    Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC    -!- FUNCTION: FORMS A VOLTAGE-INDEPENDENT POTASSIUM CHANNEL ACTIVATED

CC    BY INTRACELLULAR CALCIUM. ACTIVATION IS FOLLOWED BY MEMBRANE

CC    HYPERPOLARIZATION. THOUGHT TO REGULATE NEURONAL EXCITABILITY BY

CC    CONTRIBUTING TO THE SLOW COMPONENT OF SYNAPTIC

CC    AFTERHYPERPOLARIZATION. THE CHANNEL IS BLOCKED BY APAMIN.

CC    -!- SUBUNIT: HETERO-OLIGOMER. THE COMPLEX IS COMPOSED OF 4 CHANNEL

CC    SUBUNITS EACH OF WHICH BINDS TO A CALMODULIN SUBUNIT WHICH

```

CC CC REGULATES THE CHANNEL ACTIVITY THROUGH CALCIUM-BINDING (BY
CC CC SIMILARITY).
CC CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC CC -!- SIMILARITY: BELONGS TO THE KCNN FAMILY OF POTASSIUM CHANNELS.
CC CC -----
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; AF239613; AAG16728.1; -.
DR DR HSSP; P70604; 1G4Y.
DR DR Genew; HGNC:6291; KCNN2.
DR DR MIM; 605879; -.
DR DR InterPro; IPR004178; CaMBD.
DR DR InterPro; IPR001622; K-channel_pore.
DR DR InterPro; IPR003931; SK_channel.
DR DR Pfam; PF02888; CaMBD; 1.
DR DR Pfam; PF03530; SK_channel; 1.
DR DR Ionic channel; Transmembrane; Ion transport; Calmodulin-binding.
FT FT TRANSMEM 138 158 SEGMENT S1 (POTENTIAL).
FT FT TRANSMEM 168 188 SEGMENT S2 (POTENTIAL).
FT FT TRANSMEM 214 234 SEGMENT S3 (POTENTIAL).
FT FT TRANSMEM 256 276 SEGMENT S4 (POTENTIAL).
FT FT TRANSMEM 305 325 SEGMENT S5 (POTENTIAL).
FT FT TRANSMEM 345 365 SEGMENT S6 (PORE-FORMING) (POTENTIAL).
FT FT TRANSMEM 374 394 SEGMENT S6 (POTENTIAL).
FT FT DOMAIN 412 488 CALMODULIN-BINDING (BY SIMILARITY).
FT FT DOMAIN 41 45 POLY-GLY.
FT FT DOMAIN 51 58 POLY-ALA.
FT FT DOMAIN 83 88 POLY-GLY.
FT FT DOMAIN 91 102 POLY-GLY.
FT FT DOMAIN 563 566 POLY-ARG.
FT FT CONFLICT 52 52 D -> AA (IN REF. 2).
FT FT CONFLICT 323 323 I -> T (IN REF. 2).
FT FT CONFLICT 530 530 Q -> R (IN REF. 2).
SQ SEQUENCE 579 AA; 63803 MW; BIDAE3513C106182 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 579;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VAAWTLKA 11
DB 324 IAAWTVRA 331
:||||:|

RESULT 10
KN2 RAT STANDARD; PRT; 580 AA.
AC P70604;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small conductance calcium-activated potassium channel protein 2 (SK2).
GN KCNN2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=96376602; PubMed=8781233;
RA Koshler M., Hirschberg B., Bond C.T., Kinzie J.M., Marrion N.V.,
RA Maylie J., Adelman J.P.;
RT "Small-conductance, calcium-activated potassium channels from
mammalian brain.";
```

```

RL Science 273:1709-1714 (1996).
RN [2]
RP INTERACTION WITH CALMODULIN.
RX MEDLINE=98445090; PubMed=9774106;
RA Xia X.M., Fakler B., Rivard A.F., Wayman G., Johnson-Pais T.,
RA Keen J.B., Ishii T., Hirschberg B., Bond C.T., Lutsenko S., Maylie J.,
RA Adelman J.P.;
RT "Mechanism of calcium gating in small-conductance calcium-activated
potassium channels."
RN Nature 395:503-507 (1998).
RP [3]
RX X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 395-490 WITH CALMODULIN.
RA MEDLINE=2123356; PubMed=11323678;
RA Schumacher M.A., Rivard A.F., Bachinger H.P., Adelman J.P.;
RT "Structure of the gating domain of a Ca2+-activated K+ channel
complexed with Ca2+/calmodulin."
RL Nature 410:1120-1124 (2001).
CC -!- FUNCTION: FORMS A VOLTAGE-INDEPENDENT POTASSIUM CHANNEL ACTIVATED
BY INTRACELLULAR CALCIUM. ACTIVATION IS FOLLOWED BY MEMBRANE
HYPERPOLARIZATION. THOUGHT TO REGULATE NEURONAL EXCITABILITY BY
CONTRIBUTING TO THE SLOW COMPONENT OF SYNAPTIC
AFTERHYPERPOLARIZATION. THE CHANNEL IS BLOCKED BY APAMIN.
CC -!- SUBUNIT: HETERO-OLIGOMER. THE COMPLEX IS COMPOSED OF 4 CHANNEL
SUBUNITS EACH OF WHICH BINDS TO A CALMODULIN SUBUNIT WHICH
REGULATES THE CHANNEL ACTIVITY THROUGH CALCIUM-BINDING.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO THE KCNN FAMILY OF POTASSIUM CHANNELS.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; U69882; AAB09563.1; -.
DR DR PDB; 1G4Y; 09-MAY-01.
DR DR InterPro; IPR004178; CaMBD.
DR DR InterPro; IPR001622; K-channel_pore.
DR DR InterPro; IPR003931; SK_channel.
DR DR Pfam; PF02888; CaMBD; 1.
DR DR Pfam; PF03530; SK_channel; 1.
DR DR Ionic channel; Transmembrane; Ion transport; Calmodulin-binding;
KW 3D-structure.
FT FT TRANSMEM 140 160 SEGMENT S1 (POTENTIAL).
FT FT TRANSMEM 169 189 SEGMENT S2 (POTENTIAL).
FT FT TRANSMEM 215 235 SEGMENT S3 (POTENTIAL).
FT FT TRANSMEM 257 277 SEGMENT S4 (POTENTIAL).
FT FT TRANSMEM 306 326 SEGMENT S5 (POTENTIAL).
FT FT TRANSMEM 346 366 SEGMENT S5 (POTENTIAL).
FT FT TRANSMEM 375 395 SEGMENT S6 (POTENTIAL).
FT FT DOMAIN 413 489 CALMODULIN-BINDING.
FT FT DOMAIN 42 46 POLY-GLY.
FT FT DOMAIN 52 56 POLY-ALA.
FT FT DOMAIN 80 85 POLY-GLY.
FT FT DOMAIN 88 103 POLY-GLY.
FT FT DOMAIN 564 567 POLY-ARG.
SQ SEQUENCE 580 AA; 63847 MW; F71E0DAF7EEFA8D4 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 580;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VAAWTLKA 11
DB 325 IAAWTVRA 332
:||||:|

RESULT 11
YK81 MYCTU
ID_YK81_MYCTU STANDARD; PRT; 146 AA.
```

```

AC Q10693;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2081c.
GN RV2081c OR MT2143 OR MTCY49.20C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Braham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73966; CAA98193.1; -.
DR EMBL; AE007063; AAK46425.1; ALT_INIT.
DR TIGR; MT2143; -.
KW Tuberculin; Rv2081c; -.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
SQ SEQUENCE 146 AA; 14238 MW; BCA294986F8C9FD2 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 146;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12
DB 11 VAIWTARAA 19
||| |||
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33530 / G-37;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,

RESULT 12
ID ISPD ANASP STANDARD; PRT; 228 AA.
AC Q8YLX9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60)
DE (4-diphosphocytidyl-2C-methyl-D-erythritol synthase) (MEP
DE cytidylyltransferase) (MCT).
GN ISPD OR ALL5167.
OC Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: Catalyzes the formation of 4-diphosphocytidyl-2C-
CC methyl-D-erythritol from CTP and 2C-methyl-D-erythritol 4-
CC phosphate (by similarity).
CC -!- CATALYTIC ACTIVITY: CTP + 2-C-methyl-D-erythritol 4-phosphate =
CC diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; third step.
CC -!- SIMILARITY: BELONGS TO THE ISPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003598; BAB76866.1; -.
DR InterPro; IPR001228; ISPD_synthase.
DR Pfam; PF01128; Ispd; 1.
DR TIGRFAMs; TIGR00453; ispd; 1.
DR PROSITE; PS01295; ISPD; 1.
KW Transferase; Nucleotidyltransferase; Isoprene biosynthesis;
KW Complete proteome.
SQ SEQUENCE 228 AA; 24882 MW; 4755BF62E126AAC CRC64;

Query Match 57.1%; Score 32; DB 1; Length 228;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12
DB 30 IIAWTLAA 38
||| |||
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33530 / G-37;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,

RESULT 13
Y024_MYCGE
ID Y024_MYCGE STANDARD; PRT; 367 AA.
AC P47270;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable GTP-binding protein MG024.
GN MG024.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33530 / G-37;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,

```



RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RL Science 270:397-403(1995).  
 CC -!- SIMILARITY: BELONGS TO THE YCHF FAMILY OF GTP-BINDING PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U39681; AAC71240.1; -.  
 DR TIGR; MG024; -.  
 DR InterPro; IPR004396; Cons hypoth92.  
 DR InterPro; IPR000765; GTP1\_OBG.  
 DR Pfam; PF01018; GTP1\_OBG; 1.  
 DR PRINTS; PR00326; GTP1\_OBG.  
 DR TIGRPFAMs; TIGR00092; Cons\_hypoth92; 1.  
 KW GTP-binding; Complete proteome.  
 FT NP\_BIND 8 15 GTP (POTENTIAL).  
 FT NP\_BIND 71 75 GTP (POTENTIAL).  
 SQ SEQUENCE 367 AA; 41198 MW; 98AFCFADDE6C9BDC CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 367;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 KXVAAWTLK 10  
 DB 291 KEVRAWTFK 299  
 RESULT 14  
 ID KPBH\_MOUSE STANDARD; PRT; 406 AA.  
 AC Q9DB30;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Phosphorylase B kinase gamma catalytic chain, testis/liver isoform  
 DE (EC 2.7.1.38) (PHK-gamma-T) (Phosphorylase kinase gamma subunit 2).  
 GN PHK2G.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -!- FUNCTION: PHOSPHORYLASE B KINASE CATALYZES THE PHOSPHORYLATION OF

CC SERINE IN CERTAIN SUBSTRATES, INCLUDING TROPONIN I. THIS ISOZYME  
 CC MAY REGULATE GLYCOGENOLYSIS IN THE TESTIS.  
 CC -!- CATALYTIC ACTIVITY: 4 ATP + 2 phosphorylase B = 4 ADP +  
 CC phosphorylase A.  
 CC -!- SUBUNIT: POLYMER OF 16 CHAINS, FOUR EACH OF ALPHA, BETA, GAMMA,  
 CC AND DELTA. ALPHA AND BETA ARE REGULATORY CHAINS, GAMMA IS THE  
 CC CATALYTIC CHAIN, AND DELTA IS CALMODULIN.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AK005277; BAB23926.1; -.  
 DR HSSP; P00518; IPHK.  
 DR MGD; MGI:1916211; Phkg2.  
 DR InterPro; IPR000719; Buk\_pkinase.  
 DR InterPro; IPR002291; Phosph\_kin\_gamma.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PRO1049; PHOSPHKINASE.  
 DR PRINTS; PRO1009; TYRKINASE.  
 DR ProDom; PD000001; Buk\_pkinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TY\_Kc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Glycogen metabolism; transferase; Serine/threonine-protein kinase;  
 KW ATP-binding; Calmodulin-binding; Multigene family.  
 FT DOMAIN 24 291 PROTEIN KINASE.  
 FT NP\_BIND 30 38 ATP (BY SIMILARITY).  
 FT BINDING 53 53 ATP (BY SIMILARITY).  
 FT ACT\_SITE 153 153 BY SIMILARITY.  
 FT DOMAIN 306 330 CALMODULIN-BINDING (DOMAIN-N)  
 FT (BY SIMILARITY).  
 FT DOMAIN 346 370 CALMODULIN-BINDING (DOMAIN-C)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 406 AA; 46519 MW; 3B03CA8D9592754C CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 406;  
 Best Local Similarity 66.7%; Pred. No. 95;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 VAAWTLKAA 12  
 DB 310 VAVWTLAA 318  
 RESULT 15  
 ID KPBH\_MOUSE STANDARD; PRT; 406 AA.  
 AC P13325;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phosphorylase B kinase gamma catalytic chain, testis/liver isoform  
 DE (EC 2.7.1.38) (PHK-gamma-T) (Phosphorylase kinase gamma subunit 2).  
 GN PHK2G.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92112855; PubMed=1370475;  
 RA Calalb M.B., Fox D.T., Hanks S.K.;  
 RT "Molecular cloning and enzymatic analysis of the rat homolog of 'PHK-

```

RT gamma T,' an isoform of phosphorylase kinase catalytic subunit.";
RL J. Biol. Chem. 267:1455-1463(1992).
CC
CC -!- FUNCTION: PHOSPHORYLASE B KINASE CATALYZES THE PHOSPHORYLATION OF
CC SERINE IN CERTAIN SUBSTRATES, INCLUDING TROPONIN I. THIS ISOZYME
CC MAY REGULATE GLYCOGENOLYSIS IN THE TESTIS.
CC -!- CATALYTIC ACTIVITY: 4 ATP + 2 phosphorylase B = 4 ADP +
CC phosphorylase A.
CC -!- SUBUNIT: POLYMER OF 16 CHAINS, FOUR EACH OF ALPHA, BETA, GAMMA,
CC AND DELTA. ALPHA AND BETA ARE REGULATORY CHAINS, GAMMA IS THE
CC CATALYTIC CHAIN, AND DELTA IS CALMODULIN.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73808; AAA41863.1; -.
DR PIR; A42034; A42034.
DR HSSP; P00518; 1PHK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002291; Phosph_kin_gamma.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01049; PHOSPHBKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Glycogen metabolism; Transferase; Serine/threonine-protein kinase;
KW ATP-binding; Calmodulin-binding; Multigene family.
FT DOMAIN 24 291 PROTEIN_KINASE.
FT NP_BIND 30 38 ATP (BY SIMILARITY).
FT BINDING 53 53 ATP (BY SIMILARITY).
FT ACT_SITE 153 153 BY SIMILARITY.
FT DOMAIN 306 330 CALMODULIN-BINDING (DOMAIN-N)
FT (BY SIMILARITY).
FT DOMAIN 346 370 CALMODULIN-BINDING (DOMAIN-C)
FT (BY SIMILARITY).
SQ SEQUENCE 406 AA; 46677 MW; 7D5DB86B8E8D5BDA CRC64;

```

Query Match 57.1%; Score 32; DB 1; Length 406;

Best Local Similarity 66.7%; Pred. No. 95;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12

Db 310 VAVWTLAA 318

Search completed: January 29, 2003, 10:45:05  
Job time : 6.90909 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:40:47 ; Search time 20.8 Seconds  
(without alignments)  
128.779 Million cell updates/sec

Title: US-09-707-738-22  
Perfect score: 56  
Sequence: 1 AKXVAAMTLKAAAX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	67.9	438	12	O39493
2	37	66.1	370	16	Q8R8U0
3	37	66.1	663	3	Q96X20
4	37	66.1	1036	17	O29316
5	36	64.3	236	10	Q9XEQ8
6	36	64.3	1626	12	Q9YK98
7	35	62.5	261	17	O8ZY99
8	35	62.5	942	16	Q9HUV7
9	35	62.5	1492	2	O68724
10	35	62.5	1545	16	Q9ZH03
11	34	60.7	1072	2	Q9AZU0
12	34	60.7	170	2	Q9ZFP6
13	34	60.7	180	16	Q92RM2
14	34	60.7	201	16	Q8UEW0
15	34	60.7	297	16	Q92YE3
16	34	60.7	302	12	Q91BD8

17	34	60.7	316	16	Q93JM7
18	34	60.7	331	5	Q8SIE0
19	34	60.7	363	16	Q9PA79
20	34	60.7	368	10	Q8W366
21	34	60.7	418	17	Q9HNH3
22	34	60.7	476	10	O04927
23	34	60.7	489	16	P96223
24	34	60.7	498	16	Q9A588
25	34	60.7	499	16	Q9A588
26	34	60.7	543	13	Q90ZP5
27	34	60.7	553	13	Q9PTS9
28	34	60.7	612	5	Q8STA9
29	34	60.7	729	9	Q9ZXA5
30	34	60.7	832	16	Q8XYF4
31	34	60.7	845	5	O01914
32	34	60.7	940	16	Q928N7
33	33	58.9	163	3	Q03418
34	33	58.9	173	12	Q8QLD1
35	33	58.9	246	4	Q9H4F3
36	33	58.9	246	4	Q96DX8
37	33	58.9	295	10	Q9FZ85
38	33	58.9	295	10	Q82796
39	33	58.9	299	16	Q8UI88
40	33	58.9	338	16	Q98EX1
41	33	58.9	339	16	Q8YE75
42	33	58.9	363	16	Q8R9F4
43	33	58.9	371	16	Q9EX41
44	33	58.9	378	10	Q9SVZ0
45	33	58.9	416	2	O52976

## ALIGNMENTS

RESULT 1  
O39493 ID O39493 PRELIMINARY; PRT; 438 AA.  
AC O39493;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Glycoprotein M.  
GN Uli0.  
OS Bovine herpesvirus type 1 (strain Jura).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=31518;  
RN [1]\_TaxID=31518;  
RP SEQUENCE FROM N.A.  
RC STRAIN=JURA;  
RX MEDLINE=95313343; PubMed=7793062;  
RA Letchworth G.J., Benes V., Lu Z., Kutish G.F., Paces V., Rock D., Vilek C., Simek C., Schwyzler M.;  
RT "Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of herpes simplex virus.";  
RL Virology 210:100-108(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JURA;  
RX MEDLINE=95313343; PubMed=7793062;  
RA Letchworth G.J., Benes V., Lu Z., Kutish G.F., Paces V., Rock D., Vilek C., Simek C., Schwyzler M.;  
RT "Complete DNA sequence of bovine herpesvirus 1.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AJ004801; CA006124.1; -;  
DR InterPro; IPR000785; Herpes\_glycop.  
DR Pfam; PF01528; Herpes\_glycop; 1.  
DR PRINTS; PR00333; HSVINTEGRLMP.  
SQ SEQUENCE 438 AA; 45517 MW; 4E7C7FA64FAAEFC7 CRC64;

Query Match 67.9%; Score 38; DB 12; Length 438;  
Best Local Similarity 77.8%; Pred. No. 57;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 VAAWTLKAA 12
DB 141 LAAWTLQAA 149

RESULT 2
ID Q8R8U0 PRELIMINARY; PRT; 370 AA.
AC Q8R8U0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Kef-type K+ transport systems, membrane components.
GN KEFB OR TTE1901.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MB4T / JCM11007;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013140; AAM25084.1; -.
KW Complete proteome.
SQ SEQUENCE 370 AA; 40695 MW; 489870521F84215A CRC64;

Query Match 66.1%; Score 37; DB 16; Length 370;
Best Local Similarity 77.8%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12
DB 102 VAGWTLQAA 110

RESULT 3
ID Q96X20 PRELIMINARY; PRT; 663 AA.
AC Q96X20;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Acid phosphatase (Fragment).
OS Laccaria fraterna.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Tricholomataceae; Laccaria.
OX NCBI_TaxID=152510;
RN [1]
RP SEQUENCE FROM N.A.
RA Christy A.A., Mahadevan A.;
RT "Mineral phosphate solubilizing gene from Laccaria fraterna.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333771; AAK57642.1; -.
DR InterPro; IPR003006; IG MHC.
DR PROSITE; PS00290; IG MHC; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 663 AA; 75665 MW; 2FF821DE981BFC81 CRC64;

Query Match 66.1%; Score 37; DB 3; Length 663;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12
DB 38 VASWTLKAA 46

```

```

RESULT 4
ID O29316 PRELIMINARY; PRT; 1036 AA.
AC O29316;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein AF0946.
GN AF0946.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001039; AAB90312.1; -.
DR TIGR; AF0946; -.
DR InterPro; IPR001680; WD40.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1036 AA; 112425 MW; CA9FCE0932F9899F CRC64;

Query Match 66.1%; Score 37; DB 17; Length 1036;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12
DB 332 ASTVGAWTVREA 343

RESULT 5
ID Q9XE08 PRELIMINARY; PRT; 236 AA.
AC Q9XE08;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 25.3 kDa protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RA Liaca V., Lou A., Young S., Messing J.;
RT "Retrotransposable elements of Sorghum bicolor.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AFL14171; AAD27569.1; -.
DR InterPro; IPR001454; Hlg_nase/hydrlase.
DR InterPro; IPR004469; SerB.
DR PFam; PF00702; Hydrolase; 1.
DR TIGRFAMs; TIGR00338; serB; 1.
KW Hypothetical protein.
SQ SEQUENCE 236 AA; 25334 MW; A5C97418009CAFFA CRC64;

```

```

Query Match          64.3%; Score 36; DB 10; Length 236;
Best Local Similarity 70.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KXVAAWTLKA 11
DB 39 KAVAEWTAKA 48

RESULT 6
Q9YK98
ID Q9YK98 PRELIMINARY; PRT; 1626 AA.
AC Q9YK98;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polyprotein.
OS Blackcurrant reversion virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC Nepovirus
OX NCBI_TaxID=65743;
RN [1]
RP SEQUENCE OF 1045-1626 FROM N.A.
RX MEDLINE=98278327; PubMed=9617765;
RA Latvala S., Susi P., Kalkkinen N., Lehto K.;
RT "Characterization of the coat protein gene of mite-transmitted
RL blackcurrant reversion associated nepovirus.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20033820; PubMed=10564755;
RA Latvala-Kilby S., Lehto K.;
RT "The complete nucleotide sequence of RNA2 of blackcurrant reversion
RL nepovirus.";
RN [3]
RP SEQUENCE FROM N.A.
RA Latvala-Kilby S., Lehto K.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020051; AAD09329.2; -.
DR InterPro; IPR005054; Nepo_coat.
DR InterPro; IPR005305; Nepo_coat_C.
DR InterPro; IPR005306; Nepo_coat_N.
DR Pfam; PF03391; Nepo_coat; 1.
DR Pfam; PF03688; Nepo_coat_C; 1.
DR Pfam; PF03689; Nepo_coat_N; 1.
DR CHAIN 1094 1626 COAT PROTEIN.
SQ SEQUENCE 1626 AA; 178804 MW; 350C5BB30938263B CRC64;

Query Match          64.3%; Score 36; DB 12; Length 1626;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKXVAWTLK 10
DB 342 AKTIATWVLK 351

RESULT 7
Q8ZY99
ID Q8ZY99 PRELIMINARY; PRT; 261 AA.
AC Q8ZY99;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein PAE0883.
GN PAE0883.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AE009791; AAL63097.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 261 AA; 29211 MW; C11EF782AFC53851 CRC64;

Query Match          62.5%; Score 35; DB 17; Length 261;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12
DB 118 AKGDPANTLRS 129

RESULT 8
Q9HUV7
ID Q9HUV7 PRELIMINARY; PRT; 942 AA.
AC Q9HUV7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable sensor/response regulator hybrid.
DR PA4856.
GN Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AE004899; AAG08241.1; -.
DR HSSP; P52934; IQMP.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_KinA.
DR InterPro; IPR004359; HIS_KIN_sig.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00072; response_reg; 2.
DR Pfam; PF00512; signal; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 2.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00448; REC; 2.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 942 AA; 103855 MW; F5EC10AE3DAD85F9 CRC64;

Query Match          62.5%; Score 35; DB 16; Length 942;
Best Local Similarity 72.7%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKXVAWTLKA 11

```

```

Db 765 AKPVAGYTLKA 775
|||||:|||||
RESULT 9
O68724 PRELIMINARY; PRT; 1492 AA.
AC O68724;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Phage lambda host specific protein J.
OS Yersinia pestis.
OG Plasmid pMT1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM;
RX MEDLINE=98422474; PubMed=9748454;
RA Hu P., Elliott J., McCready P., Skowronski E., Garnes J.,
RA Kobayashi A., Brubaker R.R., Garcia E.;
RT "Structural organization of virulence-associated plasmids of Yersinia
RT pestis.";
RL J. Bacteriol. 180:5192-5202(1998).
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR EMBL; AF053947; AAC13222.1; -.
DR InterPro; IPR003961; FN_III.
DR Pfam; PF00041; fn3; 2.
KW Hydrolase; Plasmid.
SQ SEQUENCE 1492 AA; 162678 MW; 3A08640BC15A9E96 CRC64;

Query Match 62.5%; Score 35; DB 2; Length 1492;
Best Local Similarity 77.8%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAAWTLK 10
|:|||||
Db 1248 KVSAAWTLK 1256
|||||
RESULT 10
Q9ZH03 PRELIMINARY; PRT; 1545 AA.
AC Q9ZH03;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Lambda host specificity protein J.
GN Y1049 OR YPMT1.06C.
OS Yersinia pestis.
OG Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM10+; PLASMID=PMT1 (PMT-1);
RX MEDLINE=99043898; PubMed=9826348;
RA Lindler L.E., Plano G.V., Burland V., Mayhew G.F., Blattner F.R.;
RT "Complete DNA sequence and detailed analysis of the Yersinia pestis
RT KIM5 plasmid encoding murine toxin and capsular antigen.";
RL Infect. Immun. 66:5731-5742(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS; PLASMID=PMT1 (PMT-1);
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,

```

```

RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR EMBL; AF074611; AAC82709.1; -.
DR EMBL; AL117211; CAB55188.1; -.
DR InterPro; IPR003961; FN_III.
DR Pfam; PF00041; fn3; 2.
KW Hydrolase; Plasmid; Complete proteome.
FT VARIANT 1514 1514 G->V
SQ SEQUENCE 1545 AA; 168767 MW; 4E3F3B528D91DECD CRC64;

Query Match 62.5%; Score 35; DB 16; Length 1545;
Best Local Similarity 77.8%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAAWTLK 10
|:|||||
Db 1289 KVSAAWTLK 1297
|||||
RESULT 11
Q9A2U0 PRELIMINARY; PRT; 102 AA.
AC Q9A2U0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CC3466.
GN CC3466.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.R.K., Ohta N., Maddock J.R.,
RA Bisen J., Heidelberg J.F., Alley M.R.K., Stephens C., Phadke N.D., Ely B.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Swinn M.L., Haft D.H.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Swinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Barry K., O.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE006005; AAK25428.1; -.
DR TIGR; CC3466; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 102 AA; 10769 MW; AC46D2B7BE8CBF14 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 102;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKXVAAWTLKAA 12
|:|||||
Db 56 ARAGAATTKQA 67
|||||
RESULT 12
Q9ZF96 PRELIMINARY; PRT; 170 AA.
AC Q9ZF96;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PhAP.
GN PhAP.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

```

```

OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11561;
RA MEDLINE=99102223; PubMed=9882674;
RX McCool G.J., Cannon M.C.;
RT "polyhydroxyalkanoate inclusion body-associated proteins and coding
region in Bacillus megaterium.";
RL J. Bacteriol. 181:585-592(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11561;
RX MEDLINE=21311749; PubMed=11418564;
RA "Phac and Phae Are Required for Polyhydroxyalkanoic Acid Synthase
Activity in Bacillus megaterium.";
RL J. Bacteriol. 183:4235-4243(2001).
DR EMBL; AF109909; AA005256.1;
SQ SEQUENCE 170 AA; 13906 MW; 210A3408424A2B4F CRC64;

Query Match 60.7%; Score 34; DB 2; Length 170;
Best Local Similarity 60.0%; Pred. NO. 1.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KXVAWTLKA 11
Db 29 KQIEQWTLKA 38

RESULT 13
Q92RM2 PRELIMINARY; PRT; 180 AA.
AC Q92RM2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable molybdenum cofactor biosynthesis protein B.
GN MOAB OR R00843 OR SMC00863.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591785; CAC45415.1;
DR InterPro; IPR001453; MoCF biosynth.
DR Pfam; PF00994; MoCF biosynth; 1.
DR PRODOM; PD002460; MoCF biosynth; 1.
DR TIGRFAMs; TIGR00177; molyb_syn; 1.
DR PROSITE; PS01078; MoCF_BIOSYNTHESIS_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 180 AA; 20013 MW; 18F140010F410163 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 180;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12
Db 63 VKAWTLDA 71

```

```

RESULT 14
Q8UEW0 PRELIMINARY; PRT; 201 AA.
ID Q8UEW0;
AC Q8UEW0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Atul644.
GN ATU1644 OR AGR_C_3031.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Serubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollan C., Allinger M., Boughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AF009121; AAL42645.1;
DR EMBL; AF008086; AAK87420.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 201 AA; 21493 MW; 140169351A2ACC3C CRC64;

Query Match 60.7%; Score 34; DB 16; Length 201;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KXVAWTLKA 11
Db 191 REIVAWTLKS 200

RESULT 15
Q92YE3 PRELIMINARY; PRT; 297 AA.
ID Q92YE3;
AC Q92YE3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable Arg12 arginase (EC 3.5.3.1).
GN ARG12 OR RA0941 OR SMA1711.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;

```

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
RT "Nucleotide sequence and predicted functions of the entire  
RT Sinorhizobium meliloti pSymA megaplasmid";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
DR EMBL; AE007280; AAK65593.1; -.  
DR InterPro; IPR000287; Arginase.  
DR Pfam; PF00491; arginase; 1.  
KW Hydrolase; Plasmid; Complete proteome.  
SQ SEQUENCE 297 AA; 31555 MW; E2D4F0F04E5A42B9 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 297;  
Best Local Similarity 50.0%; Pred. No. 2e+02; 3; Indels 0; Gaps 0;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKXVAWTLKAA 12  
| : ||| : |  
Db 64 ARAIAAWSPKLA 75

Search completed: January 29, 2003, 10:46:42  
Job time : 22.8 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:37 ; Search time 10.1818 Seconds  
(without alignments)  
40.457 Million cell updates/sec

Title: US-09-707-738-24

Perfect score: 35

Sequence: 1 XXXXXWTLXXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_MA.\*

1: /cgn2\_6/prodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	71.4	4	1	US-08-305-871A-17
2	25	71.4	4	4	US-08-788-822A-1
3	25	71.4	12	4	US-08-788-822A-21
4	25	71.4	13	1	US-08-305-871A-22
5	25	71.4	13	1	US-08-305-871A-23
6	25	71.4	13	4	US-08-788-822A-26
7	25	71.4	13	4	US-08-788-822A-27
8	25	71.4	14	4	US-08-788-822A-30
9	25	71.4	18	2	US-08-652-369A-1
10	25	71.4	23	4	US-07-741-453A-25
11	25	71.4	25	3	US-09-075-257A-17
12	25	71.4	25	4	US-09-534-639-17
13	25	71.4	28	3	US-09-075-257A-15
14	25	71.4	28	3	US-09-075-257A-16
15	25	71.4	28	4	US-09-534-639-15
16	25	71.4	28	4	US-09-534-639-16
17	25	71.4	86	4	US-08-858-700B-471
18	25	71.4	90	1	US-07-956-700B-85
19	25	71.4	90	1	US-08-476-537-85
20	25	71.4	90	1	US-08-485-607-85
21	25	71.4	90	2	US-08-475-879-85
22	25	71.4	90	4	US-08-433-043B-85
23	25	71.4	97	2	US-08-585-585A-7
24	25	71.4	97	2	US-08-249-037C-7
25	25	71.4	97	2	US-08-788-622B-7
26	25	71.4	97	3	US-08-788-621B-7
27	25	71.4	152	4	US-08-936-165A-264

28	25	71.4	169	4	US-09-134-001C-2948	Sequence 2948, Ap
29	25	71.4	170	4	US-09-134-001C-3512	Sequence 3512, Ap
30	25	71.4	173	4	US-09-134-001C-2944	Sequence 2944, Ap
31	25	71.4	221	4	US-09-247-373B-54	Sequence 54, Appl
32	25	71.4	225	4	US-09-071-035-376	Sequence 376, App
33	25	71.4	252	4	US-09-071-035-374	Sequence 374, App
34	25	71.4	277	4	US-07-741-453A-58	Sequence 58, Appl
35	25	71.4	352	3	US-08-482-577B-2	Sequence 2, Appli
36	25	71.4	352	3	US-08-289-222E-4	Sequence 4, Appli
37	25	71.4	352	4	US-09-218-176-2	Sequence 2, Appli
38	25	71.4	352	4	US-09-054-568B-4	Sequence 4, Appli
39	25	71.4	365	1	US-08-447-500-4	Sequence 4, Appli
40	25	71.4	365	1	US-08-454-097-4	Sequence 4, Appli
41	25	71.4	365	1	US-08-453-866-4	Sequence 4, Appli
42	25	71.4	365	3	US-09-100-664A-13	Sequence 13, Appl
43	25	71.4	365	3	US-08-185-359-4	Sequence 4, Appli
44	25	71.4	365	4	US-09-335-983-13	Sequence 13, Appl
45	25	71.4	391	4	US-07-741-453A-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-08-305-871A-17

; Sequence 17, Application US/08305871A

; Patent No. 5736142

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Gaeta, Federico

; APPLICANT: Grey, Howard M.

; APPLICANT: Sidney, John

; APPLICANT: Alexander, Jeffrey L.

; TITLE OF INVENTION: Alteration of Immune Response Using Pan

; TITLE OF INVENTION: DR-Binding Peptides

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/305,871A

; FILING DATE: 14-SEP-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/121,101

; FILING DATE: 14-SEP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 14137-0062-10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-305-871A-17

Query Match 71.4%; Score 25; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 2e+05;

```
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 WTLK 10
Db 1 WTLK 4

RESULT 2
US-08-788-822A-1
; Sequence 1, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = alaninamide"
;
US-08-788-822A-21
Query Match 71.4%; Score 25; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 WTLK 10
Db 1 WTLK 4

Query Match 71.4%; Score 25; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 WTLK 10
Db 1 WTLK 4

RESULT 3
US-08-788-822A-21
; Sequence 21, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = alaninamide"
;
US-08-788-822A-21
Query Match 71.4%; Score 25; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 WTLK 10
Db 7 WTLK 10

RESULT 4
US-08-305-871A-22
; Sequence 22, Application US/08305871A
; Patent No. 5736142
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Gaeta, Federico
; APPLICANT: Grey, Howard M.
; APPLICANT: Sidney, John
; APPLICANT: Alexander, Jeffrey L.
; TITLE OF INVENTION: Alteration of Immune Response Using Pan
; TITLE OF INVENTION: DR-Binding Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```
;
;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; tyrosine or phenylalanine."
;
; US-08-305-871A-22
;
; Query Match 71.4%; Score 25; DB 1; Length 13;
; Best Local Similarity 100.0%; Pred. No. 33;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 7 WTLK 10
; Db 7 WTLK 10
;
; RESULT 5
; US-08-305-871A-23
; Sequence 23, Application US/08305871A
; Patent No. 5736142
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Gaeta, Federico
; APPLICANT: Grey, Howard M.
; APPLICANT: Sidney, John
; APPLICANT: Alexander, Jeffrey L.
; TITLE OF INVENTION: Alteration of Immune Response Using Pan
; TITLE OF INVENTION: DR-Binding Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

```
;
;
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; tyrosine or phenylalanine."
;
; US-08-305-871A-23
;
; Query Match 71.4%; Score 25; DB 1; Length 13;
; Best Local Similarity 100.0%; Pred. No. 33;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 7 WTLK 10
; Db 7 WTLK 10
;
; RESULT 6
; US-08-788-822A-26
; Sequence 26, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"
US-08-788-822A-26
Query Match 71.4%; Score 25; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 7 WTLK 10
;
RESULT 7
US-08-788-822A-27
; Sequence 27, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: Defrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine, Tyr,"
; OTHER INFORMATION: Phe or conservative substitutions
; OTHER INFORMATION: therefor"
;
; NAME/KEY: Modified-site
; LOCATION: 7..8
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa may be present or absent"
US-08-788-822A-30
Query Match 71.4%; Score 25; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 7 WTLK 10
;
RESULT 8
US-08-788-822A-30
; Sequence 30, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: Defrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine, Tyr,"
; OTHER INFORMATION: Phe or conservative substitutions
; OTHER INFORMATION: therefor"
;
; NAME/KEY: Modified-site
; LOCATION: 7..8
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "Xaa may be present or absent"
US-08-788-822A-30
Query Match 71.4%; Score 25; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 9 WTLK 12
;
RESULT 9
US-08-652-369A-1
; Sequence 1, Application US/08652369A
; Patent No. 5861268
; GENERAL INFORMATION:
```

```
QY 7 WTLK 10
Db 7 WTLK 10
```

; APPLICANT: Dean G. Tang, Kenneth V. Horn  
 ; TITLE OF INVENTION: Induction of Tumor  
 ; TITLE OF INVENTION: Cell Apoptosis With Chemical Inhibitors  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ian C. McLeod  
 ; STREET: 2190 Commons Parkway  
 ; CITY: Okemos  
 ; STATE: Michigan  
 ; COUNTRY: USA  
 ; ZIP: 48864

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette 5.25 inch,  
 ; MEDIUM TYPE: 360 kb storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS (version  
 ; OPERATING SYSTEM: 3.3)  
 ; SOFTWARE: wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,369A  
 ; FILING DATE:

; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Ian C. McLeod

; REGISTRATION NUMBER: 20,931

; REFERENCE/DOCKET NUMBER: Biomide 4.1-11

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (517) 347-4100

; TELEFAX: (517) 347-4103

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 18

; TYPE: Amino Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; MOLECULE TYPE:

; DESCRIPTION: DNA primer for segment

; DESCRIPTION: of 12-lipoxygenase.

; HYPOTHETICAL: No

; ANTI-SENSE: No

; ORIGINAL SOURCE:

; ORGANISM: human

; IMMEDIATE SOURCE:

; LIBRARY: Genomic

; US-08-652-369A-1

Query Match 71.4%; Score 25; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10

Db 1 WTLK 4

RESULT 10

US-07-741-453A-25

; Sequence 25, Application US/07741453A

; Patent No. 6228597

; GENERAL INFORMATION:

; APPLICANT: PARMENTIER, MARC

; APPLICANT: LIBERT, FREDERIC

; APPLICANT: DUMONT, JACQUES

; APPLICANT: VASSART, GILBERT

; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR

; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS

; AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES

; NUMBER OF SEQUENCES: 62

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN DARB & CUSHMAN  
 ; STREET: 1615 L STREET, N.W.  
 ; CITY: WASHINGTON, D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/741,453A  
 ; FILING DATE: 19911015

; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: KOKULIS, PAUL N.

; REGISTRATION NUMBER: 16773

; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 861-3000

; TELEFAX: (202) 822-0944

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 23 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-07-741-453A-25

Query Match 71.4%; Score 25; DB 4; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10

Db 4 WTLK 7

RESULT 11

US-09-075-257A-17

; Sequence 17, Application US/09075257A

; Patent No. 6074645

; GENERAL INFORMATION:

; APPLICANT: DIAMOND, DON JEFFREY

; APPLICANT: YORK, JOANNE

; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES

; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BART G. NEWMAN

; STREET: 555 13TH STREET, NW SUITE 701E

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/075,257A

; FILING DATE: 11-MAY-1998

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/021,298

; FILING DATE: 10-FEB-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/950,064

; FILING DATE: 14-OCT-1997

; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/747,488  
FILING DATE: 12-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: NEWLAND, BART G  
REGISTRATION NUMBER: 31,282  
REFERENCE/DOCKET NUMBER: 1954-112CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 6  
OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-17

Query Match 71.4%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 10 WTLK 13

RESULT 12  
US-09-534-639-17  
Sequence 17, Application US/09534639  
Patent No. 6251399  
GENERAL INFORMATION:  
APPLICANT: Diamond, Don J  
APPLICANT: York, Joanne  
TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPIOTOPES OF HUMAN  
FILE REFERENCE: 1954-343  
CURRENT APPLICATION NUMBER: US/09/534,639  
CURRENT FILING DATE: 2000-03-27  
EARLIER APPLICATION NUMBER: 09/075,257  
EARLIER FILING DATE: 1998-05-11  
EARLIER APPLICATION NUMBER: 09/021,298  
EARLIER FILING DATE: 1998-02-10  
EARLIER APPLICATION NUMBER: 08/950,064  
EARLIER FILING DATE: 1997-10-14  
EARLIER APPLICATION NUMBER: 08/747,488  
EARLIER FILING DATE: 1996-11-12  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (6)  
OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Human  
OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-17

Query Match 71.4%; Score 25; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||

Db 10 WTLK 13  
|||||

RESULT 13  
US-09-075-257A-15  
Sequence 15, Application US/09075257A  
Patent No. 6074645  
GENERAL INFORMATION:  
APPLICANT: DIAMOND, DON JEFFREY  
APPLICANT: YORK, JOANNE  
TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPIOTOPES  
TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BART G. NEWLAND  
STREET: 555 13TH STREET, NW SUITE 701E  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,257A  
FILING DATE: 11-MAY-1998  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/021,298  
FILING DATE: 10-FEB-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/950,064  
FILING DATE: 14-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/747,488  
FILING DATE: 12-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: NEWLAND, BART G  
REGISTRATION NUMBER: 31,282  
REFERENCE/DOCKET NUMBER: 1954-112CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 6  
OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-15

Query Match 71.4%; Score 25; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 10 WTLK 13

RESULT 14  
US-09-075-257A-16  
Sequence 16, Application US/09075257A

\*  
; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 6  
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine  
; OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-16

Query Match 71.4%; Score 25; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 10 WTLK 13

RESULT 15  
US-09-534-639-15  
; Sequence 15, Application US/09534639  
; Patent No. 6251399  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J  
; APPLICANT: York, Joanne  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
; TITLE OF INVENTION: CYTOMEGALOVIRUS  
; FILE REFERENCE: 1954-343

; CURRENT APPLICATION NUMBER: US/09/534,639  
; CURRENT FILING DATE: 2000-03-27  
; EARLIER APPLICATION NUMBER: 09/075,257  
; EARLIER FILING DATE: 1998-05-11  
; EARLIER APPLICATION NUMBER: 09/021,298  
; EARLIER FILING DATE: 1998-02-10  
; EARLIER APPLICATION NUMBER: 08/950,064  
; EARLIER FILING DATE: 1997-10-14  
; EARLIER APPLICATION NUMBER: 08/747,488  
; EARLIER FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Human  
; OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-15

Query Match 71.4%; Score 25; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 10 WTLK 13

Search completed: January 29, 2003, 10:48:22  
Job time : 10.1818 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:07 ; Search time 12.5455 Seconds  
(without alignments)  
114.943 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXXWTLXXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	69.4	14	2	leukotriene B-4 12
2	25	69.4	70	2	hypothetical prote
3	25	69.4	81	2	high potential iro
4	25	69.4	85	1	high potential iro
5	25	69.4	85	1	high potential iro
6	25	69.4	90	2	hypothetical prote
7	25	69.4	96	2	hypothetical prote
8	25	69.4	104	2	hypothetical prote
9	25	69.4	122	2	translation repres
10	25	69.4	127	2	hypothetical prote
11	25	69.4	134	2	hypothetical prote
12	25	69.4	137	2	hypothetical prote
13	25	69.4	152	2	partial transposas
14	25	69.4	152	2	hypothetical prote
15	25	69.4	159	2	hypothetical prote
16	25	69.4	160	2	hypothetical prote
17	25	69.4	163	2	hypothetical prote
18	25	69.4	164	2	conserved hypotet
19	25	69.4	172	2	hypothetical prote
20	25	69.4	173	2	transposase (21) B
21	25	69.4	175	2	hypothetical prote
22	25	69.4	177	2	ubiquitin-protein
23	25	69.4	180	2	hypothetical prote
24	25	69.4	187	2	hypothetical prote
25	25	69.4	201	2	surfactin biosynth
26	25	69.4	201	2	31K outer-membrane
27	25	69.4	201	2	hypothetical prote
28	25	69.4	201	2	conserved hypotet
29	25	69.4	203	2	conserved hypotet

30 25 69.4 207 2 C70029  
31 25 69.4 211 2 H69539  
32 25 69.4 212 1 W4WLRB  
33 25 69.4 217 2 AF1898  
34 25 69.4 218 2 D71693  
35 25 69.4 219 2 AC3615  
36 25 69.4 238 2 I67638  
37 25 69.4 239 2 I53518  
38 25 69.4 240 2 T45814  
39 25 69.4 249 2 T23077  
40 25 69.4 252 2 D87397  
41 25 69.4 263 2 S06330  
42 25 69.4 264 2 B95259  
43 25 69.4 264 2 E98124  
44 25 69.4 266 2 H98208  
45 25 69.4 266 2 A13077

#### ALIGNMENTS

##### RESULT 1

A47421  
leukotriene B-4 12-hydroxydehydrogenase (EC 1.1.1.-) - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 19-May-1995 #sequence\_revision 19-May-1995 #text\_change 17-Mar-1999  
C;Accession: A47421  
R;Yokomizo, T.; Izumi, T.; Takahashi, T.; Kaeama, T.; Kobayashi, Y.; Sato, F.; Taketani, J. Biol. Chem. 268, 18128-18135, 1993  
A;Title: Enzymatic inactivation of leukotriene B-4 by a novel enzyme found in the porcine  
A;Reference number: A47421; MUID:93352633; PMID:8394361  
A;Accession: A47421  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-14 <YOK>  
C;Keywords: oxidoreductase

Query Match 69.4%; Score 25; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 6 WTLK 9

##### RESULT 2

AB1984  
hypothetical protein asl1421 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp.  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C;Accession: AB1984  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AB1984  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-70 <KUP>  
A;Cross-references: GB:BA000019; PIDN:BA073378.1; PID:gl7130769; GSPDB:GN00179  
C;Genetics:  
A;Experimental source: strain PCC 7120  
A;Gene: asl1421

Query Match 69.4%; Score 25; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 6 WTLK 9

Db 17 WTLK 20

RESULT 3

S35586

high potential iron-sulfur protein - Chromatium tepidum

C:Species: Chromatium tepidum

C:Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 18-Jun-1999

R:Moulis, J.M.; Scherrer, N.; Gagnon, J.; Forest, E.; Petillot, Y.; Garcia, D.

Arch. Biochem. Biophys. 305, 186-192, 1993

A:Title: Primary structure of Chromatium tepidum high-potential iron-sulfur protein in x

A:Reference number: S35586; PMID:93343628; PMID:6393645

A:Accession: S35586

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-81 <MOU>

C:Superfamily: high potential iron-sulfur protein

C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein

F:43,46,61,75/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 69.4%; Score 25; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 78 WTLK 81

RESULT 4

IKHREV

high potential iron-sulfur protein [validated] - Chromatium vinosum

N:Alternate names: HiPiP

C:Species: Chromatium vinosum

C:Date: 24-Apr-1984 #sequence\_revision 24-Oct-1997 #text\_change 15-Sep-2000

C:Accession: A92330; A92143; A00263

R:Tedro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.

J. Biol. Chem. 256, 731-735, 1981

A:Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur

A:Reference number: A92330; PMID:81094036; PMID:7451471

A:Accession: A92330

A:Molecule type: protein

A:Residues: 1-73, 'D', 75-85 <TED>

R:Due, K.; Tedro, S.; Bartsch, R.G.

J. Biol. Chem. 248, 7318-7331, 1973

A:Title: The complete amino acid sequence of Chromatium high potential iron sulfur prote

A:Reference number: A92143; PMID:74012043; PMID:4745771

A:Accession: A92143

A:Molecule type: protein

A:Residues: 1-10, 'N', 12-44, 'D', 46-85 <DUS>

A:Experimental source: strain D

R:Banci, L.; Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Sompongpiwat, P.

submitted to the Brookhaven Protein Data Bank, January 1995

A:Reference number: A65814; PDB:1HRQ

C:Contents: annotation; conformation by (1)H-NMR, reduced form, residues 1-85

R:Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Sompongpiwat, P.

submitted to the Brookhaven Protein Data Bank, December 1995

A:Reference number: A66207; PDB:1NEH

C:Contents: annotation; conformation by (1)H-NMR, oxidized form, residues 1-85

R:Backes, G.; Mino, Y.; Loehr, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman

J. Am. Chem. Soc. 113, 2055-2064, 1991

A:Title: The environment of Fe4S4 clusters in ferredoxins and high-potential iron protei

A:Reference number: A44688

C:Contents: annotation; X-ray crystallography, 2.0 angstroms

A:Note: assignment of Raman spectra frequencies and hydrogen bonds around the iron-sulfu

R:Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Xuong, N.H.; Alden, R.A.; Bartsch, R.G.

J. Biol. Chem. 249, 4212-4225, 1974

A:Title: Two-angstrom crystal structure of oxidized Chromatium high potential iron prote

A:Reference number: A92153; PMID:74309824; PMID:4855287

C:Contents: annotation; X-ray crystallography, 2.0 angstroms

R:Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Alden, R.A.

J. Biol. Chem. 249, 6339-6346, 1974

A:Title: Comparison of oxidation-reduction site geometries in oxidized and reduced Chroma

A:Reference number: A92161; PMID:75019502; PMID:4417854

A:Contents: annotation; X-ray crystallography

A:Note: structures of the oxidized and reduced forms are compared with each other and wit

C:Superfamily: high potential iron-sulfur protein

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F:43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental

Query Match 69.4%; Score 25; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 80 WTLK 83

RESULT 5

IHTFER

high potential iron-sulfur protein - Thiocapsa roseopersicina (tentative sequence)

C:Species: Thiocapsa roseopersicina

C:Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 31-Mar-2000

C:Accession: A00264

R:Tedro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.

J. Biol. Chem. 256, 731-735, 1981

A:Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur

A:Reference number: A92330; PMID:81094036; PMID:7451471

A:Accession: A00264

A:Molecule type: protein

A:Residues: 1-85 <TED>

C:Superfamily: high potential iron-sulfur protein

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F:43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 69.4%; Score 25; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 80 WTLK 83

RESULT 6

D97355

hypothetical protein CAC3711 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: D97355

R:Nolling, J.; Braton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; PMID:21359325; PMID:21359325

A:Accession: D97355

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <KUR>

C:Cross-references: GB:AE001437; PIDN:AAK81631.1; PID:gi15026816; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3711

Query Match 69.4%; Score 25; DB 2; Length 90;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 46 WTLK 49

RESULT 7

## S23863

hypothetical protein 2 - garlic mosaic virus  
 C:Species: garlic mosaic virus, GarMV  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
 C:Accession: S23863  
 R:Choi, J.N.; Choi, Y.H.; Choi, Y.D.; Lee, J.S.  
 A:Description: Nucleotide sequence of a cDNA for garlic mosaic virus.  
 A:Reference number: S23862  
 A:Accession: S23863  
 A:Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 1-36 <CRO>  
 A:Cross-references: EMBL:X67134; NID:g59391; PIDN:CAA47614.1; PID:g59393

Query Match 69.4%; Score 25; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||

Db 7 WTLK 10  
 |||||

## RESULT 8

A41384  
 nicotinic acetylcholine receptor alpha chain - Chinese cobra (fragment)  
 C:Species: Naja naja atra (Chinese cobra)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: A41384  
 R:Neumann, D.; Barchan, D.; Horowitz, M.; Kochva, E.; Fuchs, S.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7255-7259, 1989  
 A:Title: Snake acetylcholine receptor: cloning of the domain containing the four extracellular loops  
 A:Reference number: A41384; MUID:89386720; PMID:2780569  
 A:Accession: A41384  
 A:Molecule type: mRNA  
 A:Residues: 1-104 <NEU>  
 A:Cross-references: GB:M26388; NID:g213370; PIDN:AAA49384.1; PID:g213371  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: neurotransmitter receptor

Query Match 69.4%; Score 25; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||

Db 58 WTLK 61  
 |||||

## RESULT 9

A45395  
 translation repressor regA - phage RB69  
 C:Species: phage RB69  
 C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: A45385  
 R:Jozwik, C.E.; Miller, E.S.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 5053-5057, 1992  
 A:Title: Regions of bacteriophage T4 and RB69 RegA translational repressor proteins that interact with the 3' noncoding region of the T4 mRNA  
 A:Reference number: A45385; MUID:92279264; PMID:1594613  
 A:Accession: A45385  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-122 <JOZ>  
 A:Cross-references: GB:M86231; NID:g215354; PIDN:AAA32295.1; PID:g215356  
 C:Genetics:  
 A:Gene: regA  
 C:Superfamily: phage T4 translation repressor  
 C:Keywords: RNA binding; translation repressor

Query Match 69.4%; Score 25; DB 2; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||

Db 112 WTLK 115  
 |||||

## RESULT 10

C97299  
 hypothetical protein CAC3248 [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: C97299  
 R:Nolling, J.; Bennett, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: C97299  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-127 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK81182.1; PID:g15026321; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3248

Query Match 69.4%; Score 25; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||

Db 103 WTLK 106  
 |||||

## RESULT 11

T12858  
 hypothetical protein yopX - Bacillus subtilis phage SPBc2  
 C:Species: Bacillus subtilis phage SPBc2  
 C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: T12858; F69918  
 R:Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.  
 submitted to the EMBL Data Library, August 1997  
 A:Description: The complete nucleotide sequence of the Bacillus subtilis SpBc2 prophage  
 A:Reference number: 217583  
 A:Accession: T12858  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-134 <LAZ>

A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025572; PIDN:AAC13067.1  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortner, C.; Bron, S.; Brouillet, S.; Brueschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallercia, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror, akuechi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F69918  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-134 <KUN>

A:Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB13991.1; PII

A:Experimental source: strain 168

C:Genetics:

A:Gene: yopX  
 C:Superfamily: Bacillus subtilis phage SPBc2 hypothetical protein yopX

Query Match 69.4%; Score 25; DB 2; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 31 WTLK 34

RESULT 12  
 G90472  
 partial transposase ISC1078 [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: G90472  
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: G90472  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-137 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3816304; PIDN:AAK43038.1; GSPDB:GN00155  
 C:Genetics:  
 A:Gene: SSO2933

Query Match 69.4%; Score 25; DB 2; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 91 WTLK 94

RESULT 13  
 F90143  
 hypothetical protein SSO0040 [imported] - Sulfolobus solfataricus transposon ISC1078  
 C:Species: Sulfolobus solfataricus  
 C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 18-Jul-2001  
 C:Accession: F90143; C90336; C90480; G90280  
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: F90143  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3813171; PIDN:AAK40405.1; GSPDB:GN00155  
 C:Accession: C90336  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3814994; PIDN:AAK41946.1; GSPDB:GN00155  
 A:Accession: C90480  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3816382; PIDN:AAK43098.1; GSPDB:GN00155  
 A:Accession: G90280  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3814460; PIDN:AAK41502.1; GSPDB:GN00155

C:Genetics:  
 A:Gene: SSO0040; SSO1750; SSO2992; SSO1264

Query Match 69.4%; Score 25; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 91 WTLK 94

RESULT 14  
 C90295  
 hypothetical protein SSO1384 [imported] - Sulfolobus solfataricus transposon ISC1078  
 C:Species: Sulfolobus solfataricus  
 C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: C90295  
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: C90295  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE006641; NID:gl3814600; PIDN:AAK41618.1; GSPDB:GN00155  
 C:Genetics:  
 A:Gene: SSO1384

Query Match 69.4%; Score 25; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 91 WTLK 94

RESULT 15  
 E72731  
 hypothetical protein APE0390 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: E72731  
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah-  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: E72731  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-159 <KAW>  
 A:Cross-references: DDBJ:AP000059; NID:gs103911; PIDN:BAA79345.1; PID:gs104029  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE0390

Query Match 69.4%; Score 25; DB 2; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 37 WTLK 40

Search completed: January 29, 2003, 10:47:36  
 Job time : 12.5455 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:30:17 ; Search time 6.81818 Seconds  
(without alignments)  
91.248 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXXWTLKXXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	25	69.4	83	1 HPIS_THETI	P80176 thermochrom
2	25	69.4	80	1 HPIS_THIRO	P00261 thiocapsa r
3	25	69.4	90	1 Y1B1_CLOAB	Q04353 clostridium
4	25	69.4	98	1 NULM_BRARE	Q9mi2 brachydanio
5	25	69.4	104	1 ACHA_NAJNA	P4143 naja naja (
6	25	69.4	114	1 ACP5_MYCGE	Q9zb79 mycoplasma
7	25	69.4	122	1 HPIS_CHRVI	P00260 chromatium
8	25	69.4	122	1 REGA_BPR69	Q01751 bacterioph
9	25	69.4	125	1 CP03_HUMAN	O95177 homo sapien
10	25	69.4	211	1 RG3A_ARCFU	O27964 archaeoglob
11	25	69.4	212	1 V84_CRPVK	P03124 cottontail
12	25	69.4	218	1 Y363_RICPR	Q9zdg6 rickettsia
13	25	69.4	233	1 MSRA_BOVIN	P34149 bos taurus
14	25	69.4	238	1 PSE2_RAT	Q63798 rattus norv
15	25	69.4	239	1 PSE2_HUMAN	Q9ul46 homo sapien
16	25	69.4	239	1 PSE2_MOUSE	P97372 mus musculu
17	25	69.4	240	1 OM31_BRUME	Q45322 bruceella me
18	25	69.4	282	1 YP08_GUITH	O78449 guillardia
19	25	69.4	294	1 UP08_ECOLI	P39173 escherichia
20	25	69.4	301	1 LEC1_RAT	P08290 rattus norv
21	25	69.4	309	1 YC30_ODOSI	P49518 odontella s
22	25	69.4	311	1 LB40_HUMAN	Q14914 homo sapien
23	25	69.4	329	1 LB4D_PIG	Q29073 sus scrofa
24	25	69.4	334	1 YEDY_ECO57	Q8xb74 escherichia
25	25	69.4	334	1 YEDY_ECOLI	P76342 escherichia
26	25	69.4	334	1 YEDY_SALTY	Q8xesi salmonella
27	25	69.4	345	1 ARGC_BACHD	Q9k8v2 bacillus ha
28	25	69.4	346	1 CANTV_GYXN	Q9pyy5 xescia c-ni
29	25	69.4	349	1 I10S_MOUSE	Q61190 mus musculu
30	25	69.4	349	1 LB4D_RABIT	Q28719 oryctolagus
31	25	69.4	349	1 MRAY_CHLPN	Q9z706 chlamydia p
32	25	69.4	349	1 Y028_BORBU	O51059 borrelia bu
33	25	69.4	352	1 IHBC_HUMAN	P55103 homo sapien

34	25	69.4	353	1 PSBD_MAIZE	P48184 zea mays (m
35	25	69.4	355	1 VAL1_ABMVW	P21947 abutilon mo
36	25	69.4	357	1 O2B2_HUMAN	Q9gsk3 homo sapien
37	25	69.4	360	1 VE2_HPV70	P50773 human papil
38	25	69.4	361	1 VAL1_PYMVU	P27258 potato yell
39	25	69.4	365	1 HHF1_SCHPO	P40235 schizosacch
40	25	69.4	370	1 VE2_HPV39	P24830 human papil
41	25	69.4	372	1 3BHS_MACMU	P27365 m 3 beta-hy
42	25	69.4	372	1 AR1B_HUMAN	O15143 homo sapien
43	25	69.4	372	1 AR1B_MOUSE	O9wv32 mus musculu
44	25	69.4	372	1 AR1B_RAT	O86556 rattus norv
45	25	69.4	373	1 NOLL_RHILO	Q52778 rhizobium 1

#### ALIGNMENTS

RESULT 1  
HPIS\_THETI  
ID HPIS\_THETI STANDARD; PRT; 83 AA.  
AC P80176;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE High potential iron-sulfur protein (HiPIP).  
GN HIP.  
OS Thermochromatium tepidum (Chromatium tepidum).  
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
OC Thermochromatium.  
OX NCBITaxID=1050;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=ATCC 43061;  
RX MEDLINE=93343628; PubMed=89393645;  
RA Moulis J.-M., Scherrer N., Gagnon J., Forest E., Petillot Y.,  
RA Garcia D.;  
RT "Primary structure of Chromatium tepidum high-potential iron-sulfur  
RT protein in relation to thermal denaturation.";  
RL Arch. Biochem. Biophys. 305:186-192(1993).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).  
RX MEDLINE=20558546; PubMed=11095707;  
RA "Nogi T., Fathir I., Kobayashi M., Nozawa T., Miki K.;  
RT "Crystal structures of photosynthetic reaction center and high-  
RT potential iron-sulfur protein from Thermochromatium tepidum:  
RT thermostability and electron transfer.";  
R Proc. Natl. Acad. Sci. U.S.A. 97:13561-13566(2000).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (0.8 ANGSTROMS).  
RX MEDLINE=22071404; PubMed=12077426;  
RA Liu L., Nogi T., Kobayashi M., Nozawa T., Miki K.;  
RT "Ultra-high-resolution structure of high-potential iron-sulfur  
RT protein from Thermochromatium tepidum.";  
RL Acta Crystallogr. D 58:1085-1091(2002).  
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +323 MV.  
CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
CC -!- MISCELLANEOUS: THIS PROTEIN IS THERMOSTABLE.  
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
CC (HIPIP) FAMILY.  
DR PIR; S35586; S35586.  
DR PDB; 1EYT; 13-DEC-00.  
DR PDB; 1IUA; 20-MAR-02.  
DR InterPro; IPR000170; Hipot\_iron-sulf.  
DR Pfam; PF01355; HiPIP; 1.  
DR PRINTS; PR00374; HIP1PFRDOXIN.  
DR PROSITE; PS00596; HiPIP; 1.  
KW Electron transport; Iron-sulfur; 4Fe-4S; 3D-structure.  
FT METAL 43 43 IRON-SULFUR (4FE-4S).  
FT METAL 46 46 IRON-SULFUR (4FE-4S).

```

FT METAL          61      75      IRON-SULFUR (4FE-4S) .
FT METAL          75      75      IRON-SULFUR (4FE-4S) .
SQ SEQUENCE      83 AA; 8786 MW; 9211684FD2C44E0A CRC64;

Query Match      69.4%; Score 25; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   |||||
Db 78 WTLK 81

RESULT 2
HPIS_THIRO
ID _HPIS_THIRO STANDARD; PRT; 85 AA.
AC P00261;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High potential iron-sulfur protein (HiPIP).
GN HIP.
OS Thiocapsa roseopersicina.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae; Thiocapsa.
OX NCBI_TaxID=1059;
RN [1]
RP MEDLINE=81094036; PubMed=7451471;
RA Tedro S.M., Meyer T.E., Bartsch R.G., Kamen M.D.;
RT "Primary structures of high potential, four-iron-sulfur ferredoxins
RT from the purple sulfur photosynthetic bacteria, Thiocapsa
RT roseopersicina and Chromatium gracile.";
RL J. Biol. Chem. 256:731-735(1981).
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HIPIP) FAMILY.
CC PIR: A00264; IPTFPR.
DR HSSP; P00260; LCKU.
DR InterPro; IPR000170; Hipot_iron_sulf.
DR Pfam; PF01355; HIPIP; 1.
DR PRINTS; PR00374; HIPIPFERDOXIN.
DR PROSITE; PS00596; HIPIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S.
FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 46 46 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 63 63 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 77 77 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 85 AA; 8889 MW; C47AD747D2218482 CRC64;

Query Match      69.4%; Score 25; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   |||||
Db 80 WTLK 83

RESULT 3
YIB1_CLOAB
ID YIB1_CLOAB STANDARD; PRT; 90 AA.
AC O04353;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein CAC3711.
GN CAC3711.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

```

---

```

OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=93273706; PubMed=8501044;
RA Sauer U., Duerre P.;
RT "Sequence and molecular characterization of a DNA region encoding a
RT small heat shock protein of Clostridium acetobutylicum.";
RL J. Bacteriol. 175:3394-3400(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65276; CAA46377.1; ALT INIT.
DR EMBL; AE007866; AAK81631.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10302 MW; DDDDE0F5C2FDD3CC CRC64;

Query Match      69.4%; Score 25; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   |||||
Db 46 WTLK 49

RESULT 4
NULM_BRARE
ID NULM_BRARE STANDARD; PRT; 98 AA.
AC Q9MIY2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB;
RX MEDLINE=21549095; PubMed=11691861;
RA Broughton R.E., Milam J.E., Roe B.A.;
RT "The complete sequence of the zebrafish (Danio rerio) mitochondrial
RT genome and evolutionary patterns in vertebrate mitochondrial DNA.";
RL Genome Res. 11:1958-1967(2001).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AC024175; AAF74305.1; -;  
 DR ZFIN; ZDB-GENE-011205-11; mtnd41.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidored4L.  
 DR Pfam; PF00420; oxidored\_g2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 98 AA; 10523 MW; 54C4B6230D23EF24 CRC64;

Query Match 69.4%; Score 25; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 ||||  
 Db 47 WTLK 50

## RESULT 5

ACHA\_NAJNA STANDARD; PRT; 104 AA.  
 AC P14143;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acetylcholine receptor protein, alpha chain (Fragment).  
 OS Naja naja (Indian cobra).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroides;  
 OC Elapidae; Elapinae; Naja.  
 OX NCBI\_TaxID=35670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89386720; PubMed=2780569;  
 RA Neumann D., Barchan D., Horowitz M., Kochva E., Fuchs S.;  
 RT "Snake acetylcholine receptor: cloning of the domain containing the  
 RT four extracellular cysteines of the alpha subunit";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7255-7259(1989).  
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA CHAINS.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M26388; AAA49384.1; -;  
 DR PIR; A41384; A41384.  
 DR InterPro; IPR001175; Neur\_channel.  
 DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
 DR PROSITE; PS00236; NEUROTROPION\_CHANNEL; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 >104  
 FT DISULFID 10 24  
 FT DISULFID 74 75  
 FT CARBOHYD 23 23  
 FT NON\_TER 104 104  
 SQ SEQUENCE 104 AA; 12194 MW; 47A39E2C9BFBA7A0 CRC64;

Query Match 69.4%; Score 25; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 6

ACPS\_MYCGE STANDARD; PRT; 114 AA.  
 AC Q9ZB79;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Holo-[acyl-carrier protein] synthase (EC 2.7.8.7) (Holo-ACP synthase)  
 DE (4'-phosphopantetheinyl transferase acps).  
 DE ACPS OR MG211.1.  
 GN ACPS OR MG211.1.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes;  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RL -!- FUNCTION: Transfers the 4'-phosphopantetheine moiety from coenzyme  
 RL A to a Ser of acyl-carrier protein (By similarity).  
 CC -!- CATALYTIC ACTIVITY: CoA + apo-[acyl-carrier protein] = adenosine  
 CC 3', 5'-bisphosphate + holo-[acyl-carrier protein].  
 CC -!- COFACTOR: Magnesium (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY. ACPS  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U39701; AAC71438.1; -;  
 DR TIGR; MG211.1; -;  
 DR InterPro; IPR002582; ACPS.  
 DR InterPro; IPR004568; Pantethn\_trn.  
 DR Pfam; PF01648; ACPS; 1.  
 DR ProDom; PD004282; ACPS; 1.  
 DR TIGRFAMs; TIGR00556; pantethn\_trn; 1.  
 KW Transferase; Lipid synthesis; Fatty acid biosynthesis; Magnesium;  
 FT COMPLETE PROTEOME.  
 FT METAL 8 8 MAGNESIUM (BY SIMILARITY).  
 FT METAL 58 58 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 114 AA; 13153 MW; 8992BBB560373E92 CRC64;

Query Match 69.4%; Score 25; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 4; Conservative 0; Mismatches 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 DB 54 WTLK 57

RESULT 7  
 ID HPIS\_CHRVI STANDARD; PRT; 122 AA.  
 AC P00260; P96753; Q9R4K3;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein precursor (HiPIP).  
 GN HiP.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochromatium.  
 OX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=97320627; PubMed=9177478;  
 RA Bruser T., Truper H.G., Dahl C.;  
 RT "Cloning and sequencing of the gene encoding the high potential iron-  
 RT sulfur protein (HiPIP) from the purple sulfur bacterium Chromatium  
 RT vinosum.";  
 RL Biochim. Biophys. Acta 1352:18-22(1997).  
 RN [2]  
 RP SEQUENCE OF 38-122.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=74012043; PubMed=4745771;  
 RA Dus K., Tedro S., Bartsch R.G.;  
 RT "The complete amino acid sequence of Chromatium high potential iron  
 RT sulfur protein.";  
 RL J. Biol. Chem. 248:7318-7331(1973).  
 RN [3]  
 RP REVISIONS TO 48; 82 AND 111.  
 RX MEDLINE=81094036; PubMed=7451471;  
 RA Tedro S.M., Meyer T.E., Bartsch R.G., Kamen M.D.;  
 RT "Primary structures of high potential, four-iron-sulfur ferredoxins  
 RT from the purple sulfur photosynthetic bacteria, Thiocapsa  
 RT roseopersicina and Chromatium gracile.";  
 RL J. Biol. Chem. 256:731-735(1981).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=74309824; PubMed=4855287;  
 RA Carter C.W. Jr., Kraut J., Freer S.T., Xuong N.H., Alden R.A.,  
 RA Bartsch R.G.;  
 RT "2-A crystal structure of oxidized Chromatium high potential iron  
 RT protein.";  
 RL J. Biol. Chem. 249:4212-4225(1974).  
 RN [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92135210; PubMed=1734968;  
 RA Nettesheim D.G., Harder S.R., Feinberg B.A., Otvos J.D.;  
 RT "Sequential resonance assignments of oxidized high-potential  
 RT iron-sulfur protein from Chromatium vinosum.";  
 RL Biochemistry 31:1234-1244(1992).  
 RN [6]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92304939; PubMed=1610810;  
 RA Gaillard J., Albrand J.-P., Moullis J.-M., Wenner D.E.;  
 RT "Sequence-specific assignments of the 1H nuclear magnetic resonance  
 RT spectra of reduced high-potential ferredoxin (HiPIP) from Chromatium  
 RT vinosum.";  
 RL Biochemistry 31:5632-5639(1992).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95118989; PubMed=7819198;

RA Banci L., Bertini I., Dikiy A., Kastrau D.H.W., Luchinat C.,  
 RA Sompornpisut P.;  
 RT "The three-dimensional solution structure of the reduced high-  
 RT potential iron-sulfur protein from Chromatium vinosum through NMR.";  
 RL Biochemistry 34:206-219(1995).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95359150; PubMed=7632685;  
 RA Bertini I., Dikiy A., Kastrau D.H., Luchinat C., Sompornpisut P.;  
 RT "Three-dimensional solution structure of the oxidized high potential  
 RT iron-sulfur protein from Chromatium vinosum through NMR. Comparative  
 RT analysis with the solution structure of the reduced species.";  
 RL Biochemistry 34:9851-9858(1995).  
 RN [9]  
 RP STRUCTURE BY NMR OF MUTANT SER-114.  
 RX MEDLINE=96216867; PubMed=8639555;  
 RA Bontrop D., Bertini I., Cappozzi F., Dikiy A., Eltis L., Luchinat C.;  
 RT "Three-dimensional structure of the reduced C77S mutant of the  
 RT Chromatium vinosum high-potential iron-sulfur protein through nuclear  
 RT magnetic resonance: comparison with the solution structure of the  
 RT wild-type protein.";  
 RL Biochemistry 35:5928-5936(1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (0.93 ANGSTROMS).  
 RX MEDLINE=20003126; PubMed=10531472;  
 RA Parisini E., Capozzi F., Lubini F., Lamzin V., Luchinat C.,  
 RA Sheldrick G.M.;  
 RT "AD initio solution and refinement of two high-potential iron protein  
 RT structures at atomic resolution.";  
 RL Acta Crystallogr. D 55:1773-1784(1999).  
 CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
 CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
 CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +360 MV.  
 CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 CC (HIPIP) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U81381; AAB48829.1; -.  
 DR PIR; A00263; IHKREV.  
 DR PDB; 1HIP; 15-APR-93.  
 DR PDB; 1HRQ; 03-JUN-95.  
 DR PDB; 1HRR; 31-JUL-95.  
 DR PDB; 1NEH; 08-MAR-96.  
 DR PDB; 1NOE; 10-JUN-96.  
 DR PDB; 1BOY; 16-DEC-98.  
 DR PDB; 1CKU; 13-MAY-99.  
 DR InterPro; IPR000170; Hipot\_ironsulph.  
 DR Pfam; PF01355; HIPIP; 1.  
 DR PRINTS; PR00374; HIPIPFERDOXIN.  
 DR PROSITE; PS00596; HIPIP; 1.  
 KW Electron transport; Iron-sulfur; 4Fe-4S; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 37  
 FT CHAIN 38 122 HIGH POTENTIAL IRON-SULFUR PROTEIN.  
 FT METAL 80 80 IRON-SULFUR (4FE-4S).  
 FT METAL 83 83 IRON-SULFUR (4FE-4S).  
 FT METAL 100 100 IRON-SULFUR (4FE-4S).  
 FT METAL 114 114 IRON-SULFUR (4FE-4S).  
 FT CONFLICT 111 111 N -> D (IN REF. 3).  
 FT TURN 41 42  
 FT STRAND 43 43  
 FT TURN 46 47



```

FT  HELIX      49      54
FT  TURN       55      55
FT  STRAND     57      57
FT  STRAND     57      57
FT  HELIX      60      62
FT  HELIX      65      68
FT  HELIX      75      77
FT  HELIX      80      82
FT  STRAND     83      83
FT  STRAND     84      85
FT  TURN       86      87
FT  TURN       88      89
FT  TURN       91      92
FT  STRAND     97      100
FT  TURN      101      102
FT  TURN      104      105
FT  STRAND    107      109
FT  TURN     110      111
FT  STRAND    113      113
FT  TURN     115      116
FT  STRAND    119      119
SQ  SEQUENCE   122 AA;  12761 MW;  ACBBAF917F32A09 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 WTLK 10
    ||||
DB  117 WTLK 120

RESULT 8
REGA BPR69
ID  REGA BPR69          STANDARD;          PRT;   122 AA.
AC  Q01751;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DE  Translation repressor protein.
GN  REGA.
OS  Bacteriophage RB69.
OC  Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC  T4-like viruses.
OX  NCBI_TaxID=12353;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92279264; PubMed=1594613;
RA  Jozwik C.E., Miller E.S.;
RT  "Regions of bacteriophage T4 and RB69 RegA translational repressor
RT  proteins that determine RNA-binding specificity.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:5053-5057(1992).
CC  -!- FUNCTION: CONTROLS THE TRANSLATION OF A NUMBER OF PROTEINS (SUCH
CC  AS REGA ITSELF, RIIIB AND AT LEAST 35 OTHERS) BY BINDING TO THEIR
CC  MRNA.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
DR  EMBL; M86231; AAA32295.1; -.
DR  HSSP; P04528; IREG.
DR  InterPro; IPR002702; Translat.reg.
DR  Pfam; PF01818; Translat.reg.1.
DR  ProDom; PD01547; Translat.reg.1.
KW  Translation regulation; Repressor; DNA-binding.
FT  DNA_BIND    15      37  H-T-H MOTIF (POTENTIAL).
FT  MUTAGEN     24      24  I->T: ACTIVITY ALTERED.
FT  MUTAGEN     25      25  A->V: 100% ACTIVITY LOSS.
FT  MUTAGEN     37      37  H->Y: 100% ACTIVITY LOSS.

```

```

FT  MUTAGEN     72      72  D->G: 100% ACTIVITY LOSS.
SQ  SEQUENCE   122 AA;  14431 MW;  C4374CE90E16D1D3 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 WTLK 10
    ||||
DB  112 WTLK 115

RESULT 9
CP03 HUMAN
ID  CP03 HUMAN          STANDARD;          PRT;   125 AA.
AC  O95177; O95176;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Protein Cl6orf3.
GN  Cl6orf3.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99009317; PubMed=9790751;
RA  Whitmore S.A., Sattasian C., Crawford J., Lower K.M., McCallum B.,
RA  Seshadri R., Corneliisse C.J., Moerland E.W., Cleton-Jansen A.-M.,
RA  Tipping A.J., Mathew C.G., Savio M., Savoia A., Verlander P.,
RA  Auerbach A.D., Van Berkel C., Pronk J.C., Doggett N.A., Callen D.F.;
RT  "Characterization and screening for mutations of the growth arrest-
RT  specific 11 (GAS11) and Cl6orf3 genes at 16q24.3 in breast cancer.";
RL  Genomics 52:325-331(1998).
CC  -!- POLYMORPHISM: THERE SEEMS TO BE TWO ALLELES OF THIS PROTEIN.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
DR  EMBL; AF050081; AAC69521.1; -.
DR  EMBL; AF050080; AAC69520.1; -.
DR  Genew; HGNC:1197; Cl6orf3.
DR  MIM; 605179; -.
KW  Repeat; Polymorphism.
FT  VARIANT      47      54  MISSING (IN SHORT ISOFORM).
FT  VARIANT      47      54  /FTID=VAR_010243.
SQ  SEQUENCE   125 AA;  12619 MW;  ACCA40DBBE0A79B4 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 81;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 WTLK 10
    ||||
DB  22 WTLK 25

RESULT 10
RS3A ARCFU
ID  RS3A ARCFU          STANDARD;          PRT;   211 AA.
AC  O27964;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  30S ribosomal protein S3ae.
GN  RPS3AE OR AF2320.
OS  Archaeoglobus fulgidus.

```

```
CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
CC Archaeoglobaceae; Archaeoglobus.
CC NCBI_TaxID=2234;
CC (1)
RN SEQUENCE FROM N.A.
RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Giehn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE S3AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF000944; AAB88936.1; -
CC TIGR; AF2320; -
CC InterPro; IPR001593; Ribosomal_S3AE.
CC Pfam; PF01015; Ribosomal_S3AE; 1.
CC ProDom; PD003035; Ribosomal_S3AE; 1.
CC PROSITE; PS01191; RIBOSOMAL_S3AE; FALSE_NEG.
CC RIBOSOMAL protein; Complete proteome.
CC SEQUENCE 211 AA; 24334 MW; 3463C7042B0F2796 CRC64;
CC -----
Query Match 69.4%; Score 25; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB 14 WTLK 17
DB -----
RESULT 11
VE4_CRPVK
ID VE4_CRPVK STANDARD; PRT; 212 AA.
AC P03124;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Probable E4 protein.
GN E4.
OS Cottontail rabbit (shope) papillomavirus (strain Kansas) (CRPV).
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=31553;
OX (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=85166175; PubMed=2984661;
RA Giri I., Danos O., Yaniv M.;
RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02708; -; NOT_ANNOTATED_CDS.
CC PIR; A03674; W4WLRB.
CC Early protein.
KW SEQUENCE 212 AA; 24104 MW; 9FB8DA3FB14ADB5B CRC64;
CC -----
Query Match 69.4%; Score 25; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB 73 WTLK 76
DB -----
RESULT 12
Y363_RICPR
ID Y363_RICPR STANDARD; PRT; 218 AA.
AC Q9ZDG6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP363.
GN RP363
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=782;
OX (1)
RN SEQUENCE FROM N.A.
RP STRAIN=Madrid E.
RX MEDLINE=99039493; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Scharitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of
RA mitochondria."
RL Nature 396:133-140(1998).
CC -!- SIMILARITY: SOME, TO R.PROWAZEKII RP364.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ235271; CAAL4822.1; -
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 218 AA; 25780 MW; 2D8BDD7DB938F954 CRC64;
CC -----
Query Match 69.4%; Score 25; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB 46 WTLK 49
DB -----
RESULT 13
MSRA_BOVIN
ID MSRA_BOVIN STANDARD; PRT; 233 AA.
AC P54149;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide methionine sulfoxide reductase (EC 1.8.4.6) (Protein-
DE methionine-S-oxide reductase) (Peptide Met(O) reductase).
DE use by non-profit institutions as long as its content is in no way
```

```

GN MSRA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Adrenal medulla;
RC MEDLINE=96312545; PubMed=8700890;
RA Moskovitz J., Weisbach H., Brot N.;
RT "Cloning the expression of a mammalian gene involved in the reduction
RT of methionine sulfoxide residues in proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2095-2099(1996).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RP MEDLINE=20519025; PubMed=11063566;
RX Lowther W.T., Brot N., Weisbach H., Matthews B.W.;
RA "Structure and mechanism of peptide methionine sulfoxide reductase, an
RT 'anti-oxidation' enzyme.";
RL Biochemistry 33:13307-13312(2000).
CC -!- FUNCTION: HAS AN IMPORTANT FUNCTION AS A REPAIR ENZYME FOR
CC PROTEINS THAT HAVE BEEN INACTIVATED BY OXIDATION. CATALYZES THE
CC REVERSIBLE OXIDATION-REDUCTION OF METHIONINE SULFOXIDE IN PROTEINS
CC TO METHIONINE.
CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC Protein L-methionine S-oxide + reduced thioredoxin.
CC -!- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37150; AAC48539.1; -.
DR PDB; 1FVA; 08-NOV-00.
DR PDB; 1FVG; 08-NOV-00.
DR InterPro; IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRPFAMs; TIGR00401; msrA; 1.
DR Oxidoreductase; 3D-structure.
KW SEQUENCE 233 AA; 25846 MW; 1F69D612723FEABA CRC64;
SQ
Query Match 69.4%; Score 25; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB 81 WTLK 84
RESULT 14
PSE2_RAT
ID PSE2_RAT STANDARD; PRT; 238 AA.
AC Q63798;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta
DE subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease
DE subunit 2) (11S regulator complex beta subunit) (REG-beta).
GN PSME2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.

```

```

RX MEDLINE=95309399; PubMed=7789512;
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,
RA Chung C.H., Shimomura N., Willy P.J., Mott J.D., Slaughter C.A.,
RA DeMartino G.N.;
RT "Primary structures of two homologous subunits of PA28, a gamma-
RT interferon-inducible protein activator of the 20S proteasome.";
RL FEBS Lett. 366:37-42(1995).
CC -!- FUNCTION: Implicated in immunoproteasome assembly and required for
CC efficient antigen processing. The PA28 activator complex enhances
CC the generation of class I binding peptides by altering the
CC cleavage pattern of the proteasome.
CC -!- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC
CC RING.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE PA28 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D45250; BAA08207.1; -.
DR HSSP; Q06323; LAVO.
DR InterPro; IPR003185; PA28_alpha.
DR InterPro; IPR003186; PA28_beta.
DR Pfam; PF02251; PA28_alpha; 1.
DR Pfam; PF02252; PA28_beta; 1.
DR Proteasome; Interferon induction.
KW SEQUENCE 238 AA; 26857 MW; AFA0F013CECE1CD3 CRC64;
SQ
Query Match 69.4%; Score 25; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB 111 WTLK 114
RESULT 15
PSE2_HUMAN
ID PSE2_HUMAN STANDARD; PRT; 239 AA.
AC Q9UL46; Q15129;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta
DE subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease
DE subunit 2) (11S regulator complex beta subunit) (REG-beta).
GN PSME2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95309399; PubMed=10199920;
RA McCusker D., Wilson M., Trowsdale J.;
RT "Organization of the genes encoding the human proteasome activators
RT PA28 alpha and PA28 beta.";
RL Immunogenetics 49:438-445(1999).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=95309399; PubMed=7789512;
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,
RA Chung C.H., Shimomura N., Willy P.J., Mott J.D., Slaughter C.A.,
RA DeMartino G.N.;
RT "Primary structures of two homologous subunits of PA28, a gamma-
RT interferon-inducible protein activator of the 20S proteasome.";
RL FEBS Lett. 366:37-42(1995).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC TISUE=Pancreea;
RL Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Implicated in immunoproteasome assembly and required for
CC efficient antigen processing. The PA28 activator complex enhances
CC the generation of class I binding peptides by altering the
CC cleavage pattern of the proteasome.
CC -!- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC
CC RING.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE PA28 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF079558; AAF02218.1; -.
DR EMBL; D45248; BAA08205.1; -.
DR EMBL; BC004368; AAH04368.1; -.
DR HSSP; Q06323; IAVO.
DR PHCI-2DPAGE; Q15129; -.
DR Genew; HGNC:9569; PSME2.
DR MIM; 602161; -.
DR InterPro; IPR003185; PA28_alpha.
DR InterPro; IPR003186; PA28_beta.
DR Pfam; PF02251; PA28_alpha.1.
DR Pfam; PF02252; PA28_beta.1.
KW Proteasome; Interferon induction.
FT CONFLICT 229 229 N -> T (IN REF. 2).
SQ SEQUENCE 239 AA; 27361 MW; 97A29563AA78A87C CRC64;

Query Match 69.4%; Score 25; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10
Db 112 WTLK 115

Search completed: January 29, 2003, 10:45:08
Job time : 7.81818 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:40:47 ; Search time 24 Seconds  
(without alignments)  
128.779 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXWTILKXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organelle.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertibrate.\*  
14: sp.unclassified.\*  
15: sp.rvirus.\*  
16: sp.bacteriap.\*  
17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	69.4	46	11 Q99LN4	Q99LN4 mus musculus
2	25	69.4	54	16 Q8VJY9	Q8VJY9 mycobacteri
3	25	69.4	60	12 Q8VAH9	Q8VAH9 white spot
4	25	69.4	64	13 Q9W7K4	Q9W7K4 najia haje.
5	25	69.4	68	2 Q8VUW6	Q8VUW6 staphylococ
6	25	69.4	70	4 Q9H1X5	Q9H1X5 homo sapien
7	25	69.4	70	16 Q8YWZ8	Q8YWZ8 anabaena sp
8	25	69.4	72	7 Q98260	Q98260 gadus morhu
9	25	69.4	89	12 Q91913	Q91913 culex nigri
10	25	69.4	92	10 Q9AR33	Q9AR33 oryza sativ
11	25	69.4	94	4 Q9HCV7	Q9HCV7 homo sapien
12	25	69.4	96	12 Q67664	Q67664 garlic late
13	25	69.4	96	15 Q9IQCI	Q9IQCI human immun
14	25	69.4	96	15 Q8QDN8	Q8QDN8 simian-huma
15	25	69.4	99	12 Q85061	Q85061 cottontail
16	25	69.4	106	4 Q96C09	Q96C09 homo sapien

17	25	69.4	121	8 Q9G5R3	Q9G5R3 celloporell
18	25	69.4	124	16 Q988A7	Q988A7 rhizobium 1
19	25	69.4	127	16 Q97E68	Q97E68 clostridium
20	25	69.4	129	2 Q93SA8	Q93SA8 staphylococ
21	25	69.4	134	9 Q64107	Q64107 bacteriopho
22	25	69.4	134	16 Q34401	Q34401 bacillus su
23	25	69.4	136	16 Q8XIG7	Q8XIG7 clostridium
24	25	69.4	137	16 Q9ACR1	Q9ACR1 streptomyc
25	25	69.4	137	17 Q97UR8	Q97UR8 sulfolobus
26	25	69.4	139	6 Q95LW6	Q95LW6 macaca fasc
27	25	69.4	139	11 Q9D8Y2	Q9D8Y2 mus musculu
28	25	69.4	143	10 Q9M673	Q9M673 cucumis mel
29	25	69.4	145	5 Q9NME6	Q9NME6 leishmania
30	25	69.4	145	16 Q8XP60	Q8XP60 clostridium
31	25	69.4	146	4 Q9H1X4	Q9H1X4 homo sapien
32	25	69.4	149	10 Q94DP9	Q94DP9 oryza sativ
33	25	69.4	150	5 Q9VQ18	Q9VQ18 drosophila
34	25	69.4	152	17 Q97YE2	Q97YE2 sulfolobus
35	25	69.4	152	17 Q97TX2	Q97TX2 sulfolobus
36	25	69.4	155	2 Q9JRE2	Q9JRE2 comamonas t
37	25	69.4	155	12 Q11842	Q11842 abutilon mo
38	25	69.4	158	12 Q56249	Q56249 tobacco rat
39	25	69.4	159	17 Q9YF51	Q9YF51 aeropyrum p
40	25	69.4	160	5 P91046	P91046 caenorhabdi
41	25	69.4	163	3 Q03418	Q03418 saccharomyc
42	25	69.4	163	10 Q9FHQ9	Q9FHQ9 arabidopsis
43	25	69.4	164	16 Q99X02	Q99X02 staphylococ
44	25	69.4	166	2 P71158	P71158 bacteroides
45	25	69.4	168	2 Q9KX75	Q9KX75 staphylococ

#### ALIGNMENTS

#### RESULT 1

Q99LN4 PRELIMINARY; PRT; 46 AA.  
AC Q99LN4;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 5.3 kDa protein.  
GN 4921519L13RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002300; AA02300.1; -.  
DR MGD; MGI:1918149; 4921519L13RIK.  
KW Hypothetical protein.  
SQ SEQUENCE 46 AA; 5337 MW; 704F20C15D045368 CRC64;

Query Match 69.4%; Score 25; DB 11; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTILK 10  
Db 37 WTILK 40

#### RESULT 2

Q8VJY9 PRELIMINARY; PRT; 54 AA.  
AC Q8VJY9

DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein MT1746.1.

GN MT1746.1.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / OSHKOSH;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RD EMBL; AE007036; AAK46016.1; -.  
RW TIGR; MT1746; -.  
KW Hypothetical protein.  
SQ SEQUENCE 54 AA; 5927 MW; 22E554F875A81BAF CRC64;  
  
Query Match 69.4%; Score 25; DB 16; Length 54;  
Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0;  
  
QY 7 WTLK 10  
Db ||||  
2 WTLK 5  
  
RESULT 3  
Q8VAH9 PRELIMINARY; PRT; 60 AA.  
AC Q8VAH9;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Wsv435 (WSSV494).  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform  
RT virus".  
RL J. Virol. 75:11811-11820(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TAIWAN;  
RX MEDLINE=20517548; PubMed=11062040;  
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
RT "Identification and characterization of a shrimp white spot syndrome  
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of  
RT cellular-type thymidine kinase and thymidylate kinase.";  
RL Virology 277:100-110(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TAIWAN;  
RX MEDLINE=21844071; PubMed=11853398;  
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
RA Lo C.F., Kou G.H.;  
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
RT spot syndrome virus and characterization of the motif important for  
RT targeting VP35 to the nuclei of transfected insect cells.";  
RL Virology 293:44-53(2002).  
RN [5]

RP SEQUENCE FROM N.A.  
RC STRAIN=TAIWAN;  
RA Lo C.F., Kou G.H.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RD EMBL; AF332093; AAL33437.1; -.  
DR EMBL; AF440570; AAL89362.1; -.  
SQ SEQUENCE 60 AA; 7091 MW; 4B4D84190256B053 CRC64;  
  
Query Match 69.4%; Score 25; DB 12; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02; Mismatches 0; Indels 0;  
Matches 4; Conservative 0;  
  
QY 7 WTLK 10  
Db ||||  
31 WTLK 34  
  
RESULT 4  
Q9W7K4 PRELIMINARY; PRT; 64 AA.  
AC Q9W7K4;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Muscle nicotinic acetylcholine receptor alpha-subunit (Fragment).  
DT ACHR-ALPHA.  
GN Naja haja.  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroidae;  
OC Elapidae; Elapinae; Naja.  
OX NCBI\_TaxID=8639;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-SKELETAL MUSCLE;  
RX MEDLINE=21396406; PubMed=11504859;  
RA Takacs Z., Wilhelm K.C., Soroca S.;  
RT "Snake alpha-Neurotoxin Binding Site on the Egyptian Cobra (Naja haja)  
RT Nicotinic Acetylcholine Receptor Is Conserved.";  
RL Mol. Biol. Evol. 18:1800-1809(2001).  
DR EMBL; AF077763; AAD43963.1; -.  
DR InterPro; IPR000566; Lipocln\_cytFABP.  
DR InterPro; IPR001175; Neur channel.  
DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
KW Receptor.  
FT NON\_TER 1 1  
FT NON\_TER 64 64  
SQ SEQUENCE 64 AA; 7639 MW; 92414F6F166DB2F4 CRC64;  
  
Query Match 69.4%; Score 25; DB 13; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0;  
Matches 4; Conservative 0;  
  
QY 7 WTLK 10  
Db ||||  
29 WTLK 32  
  
RESULT 5  
Q8VUW6 PRELIMINARY; PRT; 68 AA.  
AC Q8VUW6;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE OF13.  
OS Staphylococcus hominis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Staphylococcus.  
OX NCBI\_TaxID=1290;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC27844;

RA Katayama Y., Ito T., Hiramatsu K.;  
 RT "Localization of staphylococcal cassette chromosome in methicillin  
 RL susceptible staphylococci."  
 DR EMBL; AB063171; BAB83492.1; -;  
 SQ SEQUENCE 68 AA; 7842 MW; 57892E6E56D470DC CRC64;

Query Match 69.4%; Score 25; DB 2; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 65 WTLK 68

# RESULT 6

Q9H1X5 PRELIMINARY; PRT; 70 AA.  
 AC Q9H1X5;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE BA16L21.1.3 (NADP-dependent leukotriene B4 12-hydroxydehydrogenase  
 DE (isoform 3)).  
 GN BA16L21.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laird G.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; ALI35787; CAC22152.1; -;  
 SQ SEQUENCE 70 AA; 7983 MW; E77FF1214CF13142 CRC64;

Query Match 69.4%; Score 25; DB 4; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

# RESULT 7

Q8YWZ8 PRELIMINARY; PRT; 70 AA.  
 AC Q8YWZ8;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Asl1421.  
 GN ASL1421  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriuguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120."  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003585; BAB73378.1; -;  
 DR InterPro; IPR001387; HTH\_3.  
 DR Pfam; PF01381; HTH\_3; 1.  
 DR SMART; SM00530; HTH\_XRE; 1.

KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 70 AA; 8009 MW; 764C85F9AC336DDA CRC64;

Query Match 69.4%; Score 25; DB 16; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 17 WTLK 20

# RESULT 8

O98260 PRELIMINARY; PRT; 72 AA.  
 AC O98260;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE MHC class I antigen (Fragment).  
 GN GAMR-UA-PA1.4.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=SPLEEN;  
 RC MEDLINE=20009526; PubMed=10541806;  
 RA Persson A.C., Stet R.J.M., Pilstroem L.;  
 RT "Characterization of MHC class I and beta2-microglobulin sequences in  
 RT Atlantic cod reveals an unusually high number of expressed class I  
 RT genes".  
 RL Immunogenetics 50:49-59(1999).  
 DR EMBL; AJ133450; CAB38040.1; -;  
 DR InterPro; IPR001039; MHC\_I.  
 DR ProDom; PD000050; MHC\_I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 72 72  
 SQ SEQUENCE 72 AA; 8264 MW; FB260EAB403169BE CRC64;

Query Match 69.4%; Score 25; DB 7; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 69 WTLK 72

# RESULT 9

Q91913 PRELIMINARY; PRT; 89 AA.  
 AC Q91913;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CUN094 hypothetical protein.  
 GN CUN094.  
 OS Culex nigripalpus baculovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
 OX NCBI\_TaxID=130556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FLORIDA1997;  
 RX MEDLINE=21488685; PubMed=11602755;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kutish G.F.;  
 RT "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus."  
 RL J. Virol. 75:11157-11165(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=FLORIDA1997;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kutish G.F.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF403738; AAK941172.1; -;  
 KW Hypochemical protein.  
 SQ SEQUENCE 89 AA; 10561 MW; 01F0A22D2C792541 CRC64;

Query Match 69.4%; Score 25; DB 12; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 10

ID Q9AR33 PRELIMINARY; PRT; 92 AA.  
 AC Q9AR33;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE P0537A05.26 protein (P0554D10.2 protein).  
 GN P0537A05.26 OR P0554D10.2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Erihartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0537A05."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0554D10."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002971; BAB39297.1; -;  
 DR EMBL; AP002869; BAB39218.1; -;  
 SQ SEQUENCE 92 AA; 10627 MW; 30D6BD50E14804DD CRC64;

Query Match 69.4%; Score 25; DB 10; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 24 WTLK 27

## RESULT 11

Q9HCV7  
 ID Q9HCV7 PRELIMINARY; PRT; 94 AA.  
 AC Q9HCV7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DJ1153D9.3 (novel protein).  
 GN DJ1153D9.3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Clark G.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL109806; CAC03742.1; -;  
 SQ SEQUENCE 94 AA; 11060 MW; EOAB7D817888A06C CRC64;

Query Match 69.4%; Score 25; DB 4; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 2 WTLK 5

## RESULT 12

ID Q67664 PRELIMINARY; PRT; 96 AA.  
 AC Q67664;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF2 protein.  
 OS Garlic latent virus E29-6.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.  
 OX NCBI\_TaxID=143616;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Choi J.N., Choi Y.H., Choi Y.D., Lee J.S.;  
 RT "Nucleotide sequence of a cDNA for garlic mosaic virus."  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X67134; CAA47614.1; -;  
 SQ SEQUENCE 96 AA; 10771 MW; A48BD84E2F7CF25A CRC64;

Query Match 69.4%; Score 25; DB 12; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

## RESULT 13

ID Q91QC1 PRELIMINARY; PRT; 96 AA.  
 AC Q91QC1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vpr protein.  
 GN VPR.  
 OC Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20339596; PubMed=10881687;  
 RA Yamada I., Iwamoto A.;  
 RL Arch. Virol. 145:1021-1027(2000).  
 DR EMBL; AB034523; BAA93970.1; -;  
 DR InterPro; IPR000012; RetroV\_Vpr/X.  
 DR Pfam; PF00522; VPR; 1.  
 DR PRINTS; PR00444; HLIVPRVPX.  
 KW AIDS.  
 SQ SEQUENCE 96 AA; 11517 MW; 0C3C8C3A34FC8800 CRC64;

Query Match 69.4%; Score 25; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 14

Q8QDN8 PRELIMINARY; PRT; 96 AA.  
 AC Q8QDN8;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vpr protein.  
 GN VPR.  
 OS Simian-Human immunodeficiency virus.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=57667;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RX MEDLINE=21851104; PubMed=11861859;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RT "Evolution of a human immunodeficiency virus type 1 variant with  
 enhanced replication in pig-tailed macaque cells by DNA shuffling.";  
 RL J. Virol. 76:2924-2935(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF465242; AAL78992.1; -;  
 SQ SEQUENCE 96 AA; 11451 MW; E41BB11BF86CDD6E CRC64;

Query Match 69.4%; Score 25; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 15

Q85061 PRELIMINARY; PRT; 99 AA.  
 AC Q85061;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Papillomavirus sylvilagi Washington B (E2) and (E4) genes  
 DE (Fragment).  
 GN E4.  
 OS Cottontail rabbit papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RA Wu X., Brandsma J.;  
 RT "Papilloma Formation by Cottontail Rabbit Papillomavirus (CRPV)  
 Requires E1 and E2 Regulatory Genes in Addition to the E6 and E7  
 Transforming Genes.";  
 RL J. Virol. 0:0-0(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RX MEDLINE=85166175; PubMed=2984661;  
 RA Giri I., Danos O., Yaniv M.;  
 RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).

DR EMBL; U09496; AAB60542.2; -;  
 FT CONFLICT 28 28 V -> L (IN REF. 2).  
 FT NON\_TER 99 99  
 SQ SEQUENCE 99 AA; 11346 MW; CE2A896101FB0BF0 CRC64;  
 Query Match 69.4%; Score 25; DB 12; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 73 WTLK 76

Search completed: January 29, 2003, 10:46:45  
 Job time : 24 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:29:22 ; Search time 31.0909 Seconds  
(without alignments)  
64.288 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXWTLKXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1\_ /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	25	69.4	4	19	AAW50120
2	25	69.4	8	16	AAAR73328
3	25	69.4	8	16	AAAR73329
4	25	69.4	11	21	AAV79540
5	25	69.4	11	23	AAEL19459
6	25	69.4	11	23	AAU09839
7	25	69.4	13	16	AAAR70249
8	25	69.4	13	16	AAAR70250
9	25	69.4	13	18	AAW22120
10	25	69.4	13	18	AAW22121

11	25	69.4	13	19	AAW50125	Pan DR binding pep
12	25	69.4	13	19	AAW50126	Pan DR binding pep
13	25	69.4	13	21	AAAB36289	Promiscuous T help
14	25	69.4	13	21	AAV52558	Universal helper T
15	25	69.4	13	22	AAAB73644	Pan-DR binding pep
16	25	69.4	13	22	AAAB99709	Pan-DR-binding pep
17	25	69.4	13	22	AAAB99710	Pan-DR-binding pep
18	25	69.4	13	22	AAAB99711	Pan-DR-binding pep
19	25	69.4	13	22	AAAG84519	Pan-DR-binding epi
20	25	69.4	13	22	AAAG88271	Pan-DR-binding epi
21	25	69.4	13	22	AAAB20154	PADRE peptide. Sy
22	25	69.4	13	22	AAAB46167	PADRE universal T
23	25	69.4	13	22	AAAB49066	PADRE T-cell epit
24	25	69.4	13	23	AAJ01954	158P1D7 related HL
25	25	69.4	13	23	AAU80293	Pan DR epitope pep
26	25	69.4	13	23	ABG34860	Pan-DR-binding epi
27	25	69.4	13	23	AAU91547	Pan-DR-binding epi
28	25	69.4	13	23	ABB94471	Pan-DR-binding epi
29	25	69.4	13	23	AAU10848	Helper CD4 peptide
30	25	69.4	15	18	AAW22123	N-terminus of NADP
31	25	69.4	15	23	AAU75230	Human platelet 12-
32	25	69.4	18	20	AAW89677	Amino acids 176-19
33	25	69.4	19	16	AAAR91051	SEQ ID NO 353 from
34	25	69.4	19	20	AAV19635	Tetanus toxoid epi
35	25	69.4	20	22	AAAB46179	Tetanus toxoid epi
36	25	69.4	20	22	AAAB46182	Human APP A-beta 1
37	25	69.4	20	22	AAAB46204	Amyloid beta/PADRE
38	25	69.4	20	22	AAAB49078	Amyloid beta/PADRE
39	25	69.4	20	22	AAU49081	Human cytomagalovi
40	25	69.4	22	23	AAU10836	Human cytomagalovi
41	25	69.4	22	23	AAU10839	Human cytomagalovi
42	25	69.4	22	23	AAU10842	Human cytomagalovi
43	25	69.4	22	23	AAU10843	Human cytomagalovi
44	25	69.4	23	23	AAU10834	Human cytomagalovi
45	25	69.4	24	22	ABB36700	Peptide #4206 enco

#### ALIGNMENTS

RESULT 1  
AAW50120  
ID AAW50120 standard; peptide; 4 AA.  
XX AC AAW50120;  
XX 30-JUN-1998 (first entry)  
DT Pan DR binding peptide (17).  
DE Pan DR binding peptide (17).  
XX Pan DR binding peptide; antigen binding site; MHC molecule;  
KW DR locus.  
XX Synthetic.  
XX US5736142-A.  
XX 07-APR-1998.  
XX 14-SEP-1994; 94US-0305871.  
XX 14-SEP-1994; 94US-0305871.  
XX 14-SEP-1993; 93US-0121101.  
XX (CYTE-) CYTEL CORP.  
PA Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
XX WPI; 1998-239154/21.  
XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
PT or inducing MHC Class II mediated activation of T cells

PS Disclosure; Columns 37-38; 29pp; English.

XX The present sequence, a pan DR binding peptide, is capable of

CC binding antigen binding sites on MHC molecules, which are encoded

CC by most of the alleles of a DR locus. The peptide can be used to

CC inhibit or induce MHC Class II mediated activation of T-cells or

CC helper T-cells, which themselves mediate a CTL response. The

CC peptide can be used in mammals, especially humans, to inhibit

CC T-cell-mediated events involved in allograft rejection, allergic

CC responses and autoimmunity and as a vaccine adjuvant for enhancing

CC an immune response against an administered immunogen. The peptide

CC can be used with other immunogens to treat, e.g. prostate cancer,

CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,

CC lymphoma, CMV and condylyloma acuminatum.

XX SQ Sequence 4 AA;

Query Match 69.4%; Score 25; DB 19; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

DB 1 WTLK 4

RESULT 2

AAR73328

ID AAR73328 standard; Peptide; 8 AA.

XX AC AAR73328;

XX DT 12-DEC-1995 (first entry)

XX DE Human TSH receptor (residues 255-262).

XX KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;

XX KW antibody; affinity; detection.

XX OS Synthetic.

XX PN JP07089991-A.

XX PD 04-APR-1995.

XX PF 28-SEP-1993; 93JP-0240853.

XX PR 28-SEP-1993; 93JP-0240853.

XX PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.

XX DR WPI; 1995-167251/22.

XX CC Peptides with affinity to human TSH (thyroid stimulating hormone)

XX CC receptor antibody are used for detection of the antibody. (See also

XX CC AAR73201-592).

XX SQ Sequence 8 AA;

Query Match 69.4%; Score 25; DB 16; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

DB 4 WTLK 7

RESULT 3

AAR73329

ID AAR73329 standard; Peptide; 8 AA.

XX AC AAR73329;

XX DT 12-DEC-1995 (first entry)

XX DE Human TSH receptor (residues 257-264).

XX KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;

XX KW antibody; affinity; detection.

XX OS Synthetic.

XX PN JP07089991-A.

XX PD 04-APR-1995.

XX PF 28-SEP-1993; 93JP-0240853.

XX PR 28-SEP-1993; 93JP-0240853.

XX PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.

XX DR WPI; 1995-167251/22.

XX CC Peptides with affinity to human TSH (thyroid stimulating hormone)

XX CC receptor antibody are used for detection of the antibody. (See also

XX CC AAR73201-592).

XX SQ Sequence 8 AA;

Query Match 69.4%; Score 25; DB 16; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

DB 2 WTLK 5

RESULT 4

AAV79540

ID AAV79540 standard; Peptide; 11 AA.

XX AC AAV79540;

XX DT 15-AUG-2000 (first entry)

XX DE Pan-DR binder peptide PADRE 965.10.

XX KW Antigen presenting cell; T-lymphocyte; T-cell; immunomodulator;

XX KW autoimmune disease; allergy; cancer; infection; graft rejection;

XX KW immunotherapy; therapy; pan-DR binding peptide; PADRE; antigen.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 2 /note= "cyclohexylalanine"

XX PN WO200023053-A2.

XX PD 27-APR-2000.

XX PF 19-OCT-1999; 99WO-US24666.

XX XX

PR 20-OCT-1998; 98US-0105018.  
 XX (ALBA/) ALBANI S.  
 XX Albani S;  
 XX WPI; 2000-339492/29.  
 DR New artificial antigen presenting cells useful for isolating and  
 XX expanding T cells, and modulating T cell responses for the treatment of  
 PT e.g. autoimmune diseases, allergies -  
 XX Example 16; Page 93; 179pp; English.  
 XX This synthetic peptide, termed PADRE 965.10, is characterised as a  
 CC pan-DR binder peptide. PADRES of comparable affinity were used as  
 CC model antigens to identify human antigen-specific T cells in an  
 CC experiment designed to demonstrate that capture of T cells by  
 CC artificial antigen presenting cells (APC) is effective in  
 CC identifying polyclonal class II restricted human T cells. The  
 CC invention is directed to artificial APC, and methods of making APC  
 CC used to isolate and expand T cell populations and to modulate  
 CC T cell response. The invention also provides novel methods for  
 CC the identification and isolation and antigen-specific T cells. The  
 CC methods provide for the construction of liposomes containing  
 CC MHC:peptide complexes, accessory molecules, co-stimulatory  
 CC molecules, adhesion molecules, and other molecules irrelevant to T  
 CC cell binding or modulation that are used in the binding of  
 CC artificial APC to solid support systems that may be used in the  
 CC retrieval and identification and antigen-specific T cells. Devices  
 CC and methods are provided for treating conditions that would benefit  
 CC from modulation of T cell response, e.g. autoimmune disorders  
 CC (especially type I diabetes mellitus, multiple sclerosis,  
 CC rheumatoid arthritis, dermatomyositis, juvenile rheumatoid  
 CC arthritis and uveitis), allergies, cancer, viral infections, and  
 CC graft rejection.  
 XX Sequence 11 AA;  
 SQ

Query Match 69.4%; Score 25; DB 21; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 |||||  
 Db 6 WTLK 9

RESULT 5  
 ID AAE19459 standard; peptide; 11 AA.  
 XX AAE19459;  
 AC AAE19459;  
 XX 31-MAY-2002 (first entry)  
 DT PADRE, pan-DR binder peptide.  
 XX Human leukocyte antigen; HLA; DR-binding peptide; therapy;  
 KW stress protein; major histocompatibility complex; MHC; antiulcer;  
 KW type I diabetes; scleroderma; heat shock protein; hsp; vitiligo;  
 KW rheumatoid arthritis; lupus erythematosus; myasthenia gravis; tumour;  
 KW ulcerative colitis; infectious disease; haemostatic; nephrotropic;  
 KW polymyositis; chronic active hepatitis; primary biliary cirrhosis;  
 KW pernicious anaemia; autoimmune thyroiditis; Sjogren's syndrome; cancer;  
 KW Grave's disease; autoimmune disease; multiple sclerosis; hepatotropic;  
 KW inflammatory bowel disease; Goodpasture's syndrome; neuroprotective;  
 KW antimicrobial; immunosuppressive; dermatological; antiinflammatory;  
 KW antiinfertility; idiopathic Addison's disease; cytostatic.  
 XX Unidentified.  
 OS  
 XX Key Location/Qualifiers  
 FH

FT Misc-difference 2 /label= Unknown  
 XX FT  
 XX WO200212286-A2.  
 PN 14-FEB-2002.  
 XX PD  
 XX 08-AUG-2001; 2001WO-US41656.  
 XX PF  
 XX 09-AUG-2000; 2000US-224104P.  
 PR 06-APR-2001; 2001US-0828574.  
 XX PR  
 XX (UYCA-) UNIV CALIFORNIA SAN DIEGO.  
 XX PA  
 XX Albani S, Prakken BJ;  
 XX WPI; 2002-227137/28.  
 XX DR  
 XX Novel human leukocyte antigen pan DR-binding peptide, useful for  
 PT treating immune mediated diseases and conditions, has a fragment of  
 PT stress protein that binds to major histocompatibility complex class II  
 PT molecules -  
 XX Example 3; Page 46; 68pp; English.  
 XX The invention relates to human leukocyte antigen (HLA) pan DR-binding  
 CC peptide comprising a fragment of a stress protein that binds to one or  
 CC more major histocompatibility complex (MHC) class II molecules. The  
 CC invention also relates to heat shock protein (hsp) peptides. The peptides  
 CC of the invention and thr immunomodulating composition comprising these  
 CC peptides are useful for modulating, treating or preventing an immune-  
 CC mediated disease in a mammalian subject e.g. human, having or at risk of  
 CC having a disease including autoimmune disease, multiple sclerosis (MS),  
 CC rheumatoid arthritis, lupus erythematosus, myasthenia gravis, type I  
 CC diabetes, scleroderma, ulcerative colitis, cancer (e.g. melanoma,  
 CC lymphoma, leukaemia, lung, liver, kidney, brain, bladder solid tumours,  
 CC retinoblastoma, sarcoma and connective tissue cancers) and infectious  
 CC diseases. The peptides of the invention are also useful for screening  
 CC peptides or analogues that modulate pathogenic immune response. These  
 CC peptides are useful for treating autoimmune diseases or disorders  
 CC including Sjogren's syndrome, polymyositis, chronic active hepatitis,  
 CC mixed connective tissue disease, primary biliary cirrhosis, pernicious  
 CC anaemia, autoimmune thyroiditis, vitiligo, idiopathic Addison's disease,  
 CC gluten-sensitive enteropathy, Grave's disease, inflammatory bowel  
 CC disease, autoimmune neutropaenia, idiopathic thrombocytopenia purpura,  
 CC pemphigus vulgaris, autoimmune infertility, Goodpasture's syndrome,  
 CC bullous pemphigoid, discoid lupus and dense deposit disease. The present  
 CC sequence is PADRE, a pan-DR binder peptide used in the exemplification  
 CC of the invention.  
 XX SQ Sequence 11 AA;  
 XX

Query Match 69.4%; Score 25; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 |||||  
 Db 6 WTLK 9

RESULT 6  
 ID AAU09839 standard; peptide; 11 AA.  
 XX AAU09839  
 XX AAU09839;  
 AC AAU09839;  
 XX 26-FEB-2002 (first entry)  
 DT Pan-DR binder peptide (PADRE).  
 XX Antigenic; antidiabetic; neuroprotective; antirheumatic; antiarthritic;  
 XX dermatological; immunosuppressive; ophthalmological; antiallergic;  
 KW

KW cytostatic; virucide; antibacterial; artificial antigen presenting cell;  
KW liposome; major histocompatibility complex; MHC; T cell; allograft;  
KW cytokine production; T-helper 2 response; transplantation therapy;  
KW autoimmune disease; type I diabetes mellitus; multiple sclerosis;  
KW rheumatoid arthritis; dermatomyositis; juvenile rheumatoid arthritis;  
KW uveitis; allergy; cancer; infection; Pan-DR binder peptide; PADRE.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 2 /label= OTHER  
FT /note= cyclohexylalanine"  
XX  
XX W0200180833-A1.  
XX  
XX 01-NOV-2001.  
XX  
XX 20-APR-2000; 2000WO-IT00161.  
XX  
XX 20-APR-2000; 2000WO-IT00161.  
XX  
XX (ALBA/) ALBANI S.  
XX  
XX Albani S;  
XX  
XX WPI; 2002-055316/07.  
XX  
XX New artificial antigen presenting cell, useful for modulating T cell  
PT response for treating allergies and cancers, comprises liposome, major  
PT histocompatibility complex, antigen and accessory molecule components -  
XX  
XX Example 16; Page 103; 195pp; English.  
XX  
XX The invention relates to an artificial antigen presenting cell (1)  
CC comprising liposome, major histocompatibility complex (MHC), antigen and  
CC accessory molecule components. (1) is used in methods for the  
CC following: (1) identifying T cells specific for an antigen of interest;  
CC (2) isolating T cells specific for an antigen of interest; (3) modulating  
CC T cell response; (4) characterising the functional state of antigen-  
CC specific T cells; (5) treating a condition in a subject which would be  
CC benefited by altering the functional pattern of cytokine production by  
CC certain antigen-specific T cells to increase T-helper (Th) 2 response  
CC and/or decrease Th1 response; (6) identifying antigen-specific T cells  
CC specific for epitopes on a graft donor's tissue likely to elicit graft  
CC versus host rejection response; and (7) treating a recipient mammal to  
CC reduce rejection of allografts in a transplantation therapy regime.  
CC Method (5) is useful for treating autoimmune disease such as type I  
CC diabetes mellitus, multiple sclerosis, rheumatoid arthritis,  
CC dermatomyositis, juvenile rheumatoid arthritis or uveitis. Alternatively  
CC it is useful for treating allergy due to allergens such as dust, animal  
CC skin bypass products, vegetables, fruits, pollen or chemicals, cancer,  
CC viral infection, bacterial infection. Addition of the accessory  
CC molecules, as well as co-stimulatory molecules, and other proteins in  
CC proper orientation in the liposomes allow for substantially improved  
CC binding association and manipulation of T cells which is very important  
CC in the identification and stimulation of antigen-specific T cells. The  
CC present sequence represents Pan-DR binder peptide (PADRE) used in  
CC the method of the invention.  
XX  
XX  
XX Sequence 11 AA;  
Query Match 69.4%; Score 25; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 WTLK 10  
Db 6 WTLK 9  
XX  
XX  
XX RESULT 7  
XX AAR70249

ID AAR70249 standard; peptide; 13 AA.  
XX  
AC AAR70249;  
XX  
DT 13-NOV-1995 (first entry)  
XX  
XX Pan DR-binding peptide activator of T cells.  
XX  
XX T cell; MHC; class II; major histocompatibility complex II;  
KW immunogen; activation.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 1 /note= "D-form residue"  
FT  
FT Modified-site 3 /label= OTHER  
FT /note= "cyclohexylalanine"  
FT Misc-difference 13 /note= "D-form residue"  
FT  
XX W09507707-A.  
XX  
XX 23-MAR-1995.  
XX  
XX 14-SEP-1994; 94WO-US10368.  
XX  
XX 14-SEP-1993; 93US-0121101.  
XX  
XX (CYTE-) CVTEL CORP.  
XX  
XX Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
XX  
XX WPI; 1995-131178/17.  
XX  
XX Inhibiting or inducing an immune response using Pan DR-binding  
PT peptide(s) - to prepare a compsn. used to treat allo:graft  
PT rejection, allergic response and auto:immunity and as a vaccine  
PT component  
XX  
XX Claim 3; Page 51; 59pp; English.  
XX  
XX Peptides capable of binding MHC class II molecules via antigen  
CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
CC AAR70249 is capable of binding MHC class II molecules causing T cell  
CC activation. This is useful as a vaccine component and may be used in  
CC conjunction with CTL peptides to enhance a CTL response.  
XX  
XX Sequence 13 AA;  
SQ  
Query Match 69.4%; Score 25; DB 16; Length 13;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 WTLK 10  
Db 7 WTLK 10  
XX  
XX  
XX RESULT 8  
XX AAR70250  
XX  
XX AAR70250 standard; peptide; 13 AA.  
XX  
XX AAR70250;  
XX  
XX 13-NOV-1995 (first entry)  
XX  
XX Pan DR-binding peptide activator of T cells.  
XX  
XX T cell; MHC; class II; major histocompatibility complex II;  
KW immunogen; activation.  
XX

OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13  
 FT Misc-difference 13 /note= "D-form residue"  
 XX  
 PN WO9507707-A.  
 XX  
 XX 23-MAR-1995.  
 PD  
 XX 14-SEP-1994; 94WO-US10368.  
 PP  
 XX 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
 XX WPI; 1995-131178/17.  
 DR  
 XX Inhibiting or inducing an immune response using Pan DR-binding  
 PT peptide(s) - to prepare a compsn. used to treat allo:graft  
 PT rejection, allergic response and auto:immunity and as a vaccine  
 PT component  
 XX  
 PS Claim 3; Page 51; 59pp; English.  
 XX  
 CC Peptides capable of binding MHC class II molecules via antigen  
 CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
 CC AAR70250 is capable of binding MHC class II molecules causing T cell  
 CC activation. This can be used as a vaccine component. The peptide can also  
 CC be used in combination with CTL peptides to enhance a CTL response.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 69.4%; Score 25; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 ||||  
 Db 7 WTLK 10  
 ||||  
 RESULT 9  
 AAW22120  
 ID AAW22120 standard; peptide; 13 AA.  
 XX  
 AC AAW22120;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 5.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Modified-site 3 /note= "cyclohexylalanine"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 XX WO9726784-A1.  
 PN  
 XX 31-JUL-1997.  
 PD  
 XX

PF 23-JAN-1997; 97WO-US01041.  
 XX  
 PR 24-JAN-1996; 96US-0010510.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 DR  
 XX Composition for eliciting immune response to non-protein determinant  
 FT - comprises the determinant and a pan-DR binding peptide, used for  
 FT prevention and treatment of tumours and infections  
 XX  
 PS Claim 23; Page 74; 87pp; English.  
 XX  
 CC This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 69.4%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 ||||  
 Db 7 WTLK 10  
 ||||  
 RESULT 10  
 AAW22121  
 ID AAW22121 standard; peptide; 13 AA.  
 XX  
 AC AAW22121;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 6.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 XX WO9726784-A1.  
 PN  
 XX 31-JUL-1997.  
 PD  
 XX 23-JAN-1997; 97WO-US01041.  
 PF  
 XX 24-JAN-1996; 96US-0010510.  
 PR  
 XX

PA (CYTE-) CYTEL CORP.  
 XX Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 XX  
 XX Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 PT prevention and treatment of tumours and infections  
 XX  
 XX Claim 23; Page 74; 87pp; English.  
 XX  
 XX This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 XX Sequence 13 AA;  
 SQ

Query Match 69.4%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

RESULT 11  
 AAW50125  
 ID AAW50125 standard; peptide; 13 AA.  
 XX  
 XX AAW50125;  
 AC  
 DT 30-JUN-1998 (first entry)  
 XX  
 XX Pan DR binding peptide (22).  
 DE  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 KW  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 3 /label= Tyr, Phe  
 FT  
 XX US5736142-A.  
 PN  
 XX 07-APR-1998.  
 PD  
 XX 14-SEP-1994; 94US-0305871.  
 PF  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 PI WPI; 1998-239154/21.  
 XX

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 39-40; 29pp; English.  
 PS  
 XX  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or  
 CC helper T-cells, which themselves mediate a CTL response. The  
 CC peptide can be used in mammals, especially humans, to inhibit  
 CC T-cell-mediated events involved in allograft rejection, allergic  
 CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
 CC an immune response against an administered immunogen. The peptide  
 CC can be used with other immunogens to treat, e.g. prostate cancer,  
 CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
 CC lymphoma, CMV and condyloma acuminatum.  
 CC  
 XX Sequence 13 AA;  
 SQ

Query Match 69.4%; Score 25; DB 19; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

RESULT 12  
 AAW50126  
 ID AAW50126 standard; peptide; 13 AA.  
 XX  
 XX AAW50126;  
 AC  
 DT 30-JUN-1998 (first entry)  
 XX  
 XX Pan DR binding peptide (23).  
 DE  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 KW  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 3 /label= Tyr, Phe  
 FT  
 XX US5736142-A.  
 PN  
 XX 07-APR-1998.  
 PD  
 XX 14-SEP-1994; 94US-0305871.  
 PF  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 PI WPI; 1998-239154/21.  
 XX

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 41-42; 29pp; English.  
 PS  
 XX  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or



CC helper T-cells, which themselves mediate a CTL response. The  
 CC peptide can be used in mammals, especially humans, to inhibit  
 CC T-cell-mediated events involved in allograft rejection, allergic  
 CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
 CC an immune response against an administered immunogen. The peptide  
 CC can be used with other immunogens to treat, e.g. prostate cancer,  
 CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
 CC lymphoma, CMV and condyloma acuminatum.

XX SQ Sequence 13 AA;

Query Match 69.4%; Score 25; DB 19; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 DB 7 WTLK 10

# RESULT 13

AAB36289  
 ID AAB36289 standard; Peptide; 13 AA.

XX AC AAB36289;

XX DT 26-FEB-2001 (first entry)

XX DE Promiscuous T helper epitope SEQ ID NO: 65.

XX KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;  
 KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX OS Synthetic.

XX PN WO200065058-A1.

XX PD 02-NOV-2000.

XX PF 19-APR-2000; 2000WO-DK00205.

XX PR 23-APR-1999; 99DK-0000552.

XX PR 06-MAY-1999; 99US-0132811.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Klyser S;

XX DR WPI; 2000-672791/65.

XX PT Down-regulating interleukin 5 (IL-5) activity in humans by  
 PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,  
 PT prophylaxis or amelioration of asthma or other chronic allergic  
 PT conditions -

XX PS Disclosure; Page 168; 172pp; English.

XX CC The present invention is concerned with methods of treating asthma,  
 CC eosinophilia, allergic rhinitis and other allergic diseases. These  
 CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
 CC proteins and their coding sequences to down-regulate IL-5 activity and  
 CC thus reduce eosinophil numbers. The allergic diseases may be treated  
 CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
 CC it is possible that they may be used in the treatment of cancer and  
 CC helminthic infections.

XX SQ Sequence 13 AA;

Query Match 69.4%; Score 25; DB 21; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10

DB 7 WTLK 10

# RESULT 14

AAY52558

ID AAY52558 standard; peptide; 13 AA.

XX AC AAY52558;

XX DT 28-FEB-2000 (first entry)

XX DE Universal helper T epitope, pan DR epitope (PADRE).

XX KW Chimeric; Ii protein; pan DR epitope; expression vector;  
 KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
 KW epitope; antigen; presentation; class I; cytosolic pathway;  
 KW endoplasmic reticulum; class II; extracellular antigen;  
 KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;  
 KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
 KW vaccine; immunity; infection; pathogen; virus; HIV; HCV; HCV;  
 KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
 KW autoimmune disease; activation; antiviral; antimalarial;  
 KW immunoprotective.

XX OS Synthetic.

XX PN WO958658-A2.

XX PD 18-NOV-1999.

XX PF 13-MAY-1999; 99WO-US10646.

XX PR 13-MAY-1998; 98US-0078904.

XX PR 15-MAY-1998; 98US-0085751.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Fikes JD, Hermanson GG, Sette A, Ishioka GV, Livingston B;  
 PI Chesnut RW;

XX DR WPI; 2000-039103/03.

XX DR N-PSDB; AAZ38679.

XX PT Expression vectors encoding major histocompatibility targeting  
 PT sequence, used as, e.g. tumor vaccines -

XX PS Claim 9; Page 80; 130pp; English.

XX CC This sequence represents a universal helper T epitope, pan DR epitope  
 CC (PADRE), DNA encoding which is used to construct fusion genes used  
 CC in exemplifications of the present invention. The invention  
 CC relates to a novel expression vector comprising a promoter operably  
 CC linked to a fusion gene encoding a major histocompatibility complex  
 CC (MHC) targeting sequence, and two or more heterologous peptide epitopes.  
 CC The MHC targeting sequence may be a class I targeting sequence, which  
 CC directs an MHC class I epitope to a cytosolic pathway or to the  
 CC endoplasmic reticulum, or an MHC class II targeting sequence, which  
 CC directs extracellular antigens to enter the endocytic pathway to be  
 CC processed into antigen peptides for presentation on MHC class II  
 CC molecules. The heterologous epitopes may comprise either helper T  
 CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and  
 CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are  
 CC useful for stimulating an immune response in vivo, as well as for use in  
 CC assaying the human immunogenicity of a human T cell peptide epitope in  
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
 CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
 CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
 CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
 CC combined with other MHC class I and class II epitopes to increase the  
 CC number of cells that are activated in response to a given antigen and  
 CC provide a broader population coverage of MHC-reactive alleles.

XX SQ Sequence 13 AA;  
 Query Match 69.4%; Score 25; DB 21; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 Db 7 WTLK 10

RESULT 15  
 AAB73644  
 ID AAB73644 standard; peptide; 13 AA.  
 XX  
 AC AAB73644;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Pan-DR binding peptide, PADRE.  
 XX  
 KW Cellular vaccine; antigen-presenting cell; APC;  
 KW cell-surface molecule density; major histocompatibility complex; MHC;  
 KW antigen-specific T-cell expansion; tumour; cancer; viral infection;  
 KW parasitic infection; Pan-DR binding peptide; PADRE.  
 XX  
 OS Unidentified.  
 XX  
 FN WO200136978-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 15-NOV-2000; 2000WO-US42213.  
 XX  
 PR 15-NOV-1999; 99US-0165428.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schultze JL, Vonderheide RH, Nadler LM, Maecker B;  
 PI Von Bergwelt-Baildon M;  
 XX  
 DR WPI; 2001-343909/36.

Determining surface density of molecules on antigen-presenting cells,  
 useful for quality control of cellular vaccines for treatment of tumors  
 and infections comprises measuring density ratios -

Disclosure; Page 21; 47pp; English.

The invention relates to a method of determining the surface density of  
 a cell-surface molecule on a primary or artificial antigen-presenting  
 cell (APC). The method comprises determining the cell surface area  
 of the APCs, determining the absolute amount of cell surface molecule on  
 the surface of the APCs, and calculating the ratio of the amount of  
 the cell surface molecule to the APC surface area as a measure of cell  
 surface molecule density. The invention also encompasses a kit for  
 determining the cell-surface density of a complex of peptide and MHC  
 (major histocompatibility complex) protein/peptide complex on an APC;  
 CC identifying an agent that increases persistence of the MHC/peptide  
 CC complex on an APC surface by culturing APCs in presence of test compound  
 CC and measuring the time of persistence relative to an untreated control;  
 CC and determining if the APC carries a therapeutically adequate amount of  
 CC peptide by determining whether the cell-surface density of the  
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The  
 CC method is used to determine if the amount of MHC/antigenic peptide  
 CC complex present on the surface of an APC is sufficient for therapeutic  
 CC use of the cells as cellular vaccines or for ex vivo expansion of  
 CC antigen-specific T cells for subsequent return to the patient,  
 CC particularly for the treatment of tumours or viral or parasitic  
 CC infections. The method can also be used to identify agents that increase  
 CC production of MHC/peptide complexes on cells, (which increases the  
 CC therapeutic potential of the cells) by treatment before and/or during

CC APC administration. The method allows identification of therapeutically  
 CC useful antigen-pulsed APC optimisation of conditions for their  
 CC production, and control of APC quality. Sequences AAB73643-AAB73648  
 CC represent peptides used in MHC binding studies and for the generation  
 CC of peptide-specific cytotoxic T-lymphocytes.

XX SQ Sequence 13 AA;

Query Match 69.4%; Score 25; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 Db 7 WTLK 10

Search completed: January 29, 2003, 10:44:35  
 Job time : 32.0909 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:46:52 ; Search time 6.54545 Seconds  
(without alignments)  
46.243 Million cell updates/sec

Title: US-09-707-738-25

Perfect score: 36

Sequence: 1 XXXXXWTLKXXXX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122236 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	69.4	11	10	US-09-756-983-7
2	25	69.4	11	10	US-09-828-574-12
3	25	69.4	13	9	US-09-949-375A-18
4	25	69.4	13	9	US-09-785-215-19
5	25	69.4	13	10	US-09-894-018-69
6	25	69.4	24	10	US-09-864-761-47626
7	25	69.4	24	10	US-09-983-019-10
8	25	69.4	27	10	US-09-864-761-40242
9	25	69.4	27	10	US-09-983-019-4
10	25	69.4	27	10	US-09-983-019-7
11	25	69.4	28	10	US-09-864-761-39731
12	25	69.4	48	10	US-09-864-761-45303
13	25	69.4	51	9	US-10-016-634A-165
14	25	69.4	55	9	US-09-949-375A-12
15	25	69.4	56	10	US-09-864-761-41804
16	25	69.4	57	10	US-09-864-761-44535
17	25	69.4	58	10	US-09-925-300-1859
18	25	69.4	59	10	US-09-925-299-822
19	25	69.4	62	10	US-09-896-578-5

20	69.4	73	10	US-09-864-761-47212	Sequence 47212, A
21	69.4	74	10	US-09-764-877-1943	Sequence 1943, Ap
22	69.4	80	10	US-09-894-018-97	Sequence 97, Appl
23	69.4	87	10	US-09-925-297-562	Sequence 562, App
24	69.4	89	10	US-09-867-550-686	Sequence 686, App
25	69.4	97	10	US-09-864-761-43929	Sequence 43929, A
26	69.4	98	10	US-09-894-018-101	Sequence 101, Appl
27	69.4	106	9	US-10-138-516-3	Sequence 3, Appli
28	69.4	106	9	US-10-146-110-5	Sequence 5, Appli
29	69.4	106	10	US-09-894-018-95	Sequence 95, Appl
30	69.4	107	10	US-09-894-018-93	Sequence 93, Appl
31	69.4	107	10	US-09-894-018-103	Sequence 103, App
32	69.4	109	10	US-09-910-150-32	Sequence 32, Appl
33	69.4	113	9	US-09-789-054A-18	Sequence 18, Appl
34	69.4	123	10	US-09-894-018-109	Sequence 109, App
35	69.4	130	10	US-09-894-018-99	Sequence 99, Appl
36	69.4	144	10	US-09-894-018-129	Sequence 129, App
37	69.4	147	10	US-09-894-018-131	Sequence 131, App
38	69.4	148	10	US-09-894-018-127	Sequence 127, App
39	69.4	152	10	US-09-939-980-264	Sequence 264, App
40	69.4	157	10	US-09-894-018-117	Sequence 117, App
41	69.4	168	10	US-09-894-018-115	Sequence 115, App
42	69.4	169	10	US-09-894-018-119	Sequence 119, App
43	69.4	170	9	US-09-479-040-3	Sequence 3, Appli
44	69.4	180	10	US-09-894-018-125	Sequence 125, App
45	69.4	183	9	US-09-883-152-19	Sequence 19, Appl

#### ALIGNMENTS

RESULT 1  
US-09-756-983-7  
; Sequence 7, Application US/09756983  
; Patent No. US20020122818A1  
; GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthesized peptide totally artificial  
; OTHER INFORMATION: Xaa in position 2 stands for cyclohexylalanine  
US-09-756-983-7

Query Match 69.4%; Score 25; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
Db 6 WTLK 9

RESULT 2  
US-09-828-574-12  
; Sequence 12, Application US/09828574  
; Patent No. US20020146759A1

```
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: ALBANI, Salvatore
; APPLICANT: PRACKEN, Berent J.
; TITLE OF INVENTION: STRESS PROTEINS AND PEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: UCSd1310-1
; CURRENT APPLICATION NUMBER: US/09/828,574
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/224,104
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is any amino acid
US-09-828-574-12

Query Match          69.4%; Score 25; DB 10; Length 11;
Best Local Similarity 100.0%; Pred.No. 17;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 6 WTLK 9

RESULT 3
US-09-949-375A-18
; Sequence 18, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.
US-09-949-375A-18

Query Match          69.4%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred.No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 7 WTLK 10

RESULT 4
US-09-785-215-19
; Sequence 19, Application US/09785215
; Publication No. US2002019157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 19
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial T-cell epitope capable of binding to a large portion
; OTHER INFORMATION: of MHC Class II molecules in a variety of animals
US-09-785-215-19

Query Match          69.4%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred.No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 7 WTLK 10

RESULT 5
US-09-894-018-69
; Sequence 69, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Dennis
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69

Query Match          69.4%; Score 25; DB 10; Length 13;
Best Local Similarity 100.0%; Pred.No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 7 WTLK 10

RESULT 6
US-09-864-761-47626
; Sequence 47626, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
```

```
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47626
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011502.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: AI033706.1, EVALUE 1.00e-08
; OTHER INFORMATION: SWISSPROT HIT: Q84291, EVALUE 4.90e+00
US-09-864-761-47626
```

```
Query Match 69.4%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10
Db 13 WTLK 16

RESULT 7
US-09-983-019-10
; Sequence 10, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (3)..(3)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-10

Query Match 69.4%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10
Db 7 WTLK 10

RESULT 8
US-09-864-761-40242
; Sequence 40242, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40242
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC025746.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-09-864-761-40242

```

```

Query Match          69.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
Db 1 WTLK 4

```

## RESULT 9

```

US-09-983-019-7
; Sequence 4, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2001-10-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (6)..(6)
; OTHER INFORMATION: x = cyclohexylalanine or phenylalanine
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-4

```

```

Query Match          69.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
Db 10 WTLK 13

```

## RESULT 10

```

US-09-983-019-7
; Sequence 7, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US

```

```

; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (17)..(17)
; OTHER INFORMATION: x = cyclohexylalanine or phenylalanine
US-09-983-019-7

```

```

Query Match          69.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
Db 21 WTLK 24

```

## RESULT 11

```

US-09-864-761-39731
; Sequence 39731, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687

```

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39731
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003681.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
US-09-864-761-39731

```

```

Query Match          69.4%; Score 25; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
DB 11 WTLK 14

```

```

RESULT 12
US-09-864-761-45303
; Sequence 45303, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45303
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC022216.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.71
; OTHER INFORMATION: SWISSPROT HIT: Q9ZER9, EVALUE 1.20e+00
US-09-864-761-45303

```

```

Query Match          69.4%; Score 25; DB 10; Length 48;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
DB 2 WTLK 5

```

```

RESULT 13
US-10-016-634A-165
; Sequence 165, Application US/10016634A
; Publication No. US20020192666A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Ghosh, Malavika
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Prot
; FILE REFERENCE: DEX-0255
; CURRENT APPLICATION NUMBER: US/10/016,634A
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,258
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-634A-165

```

```

Query Match          69.4%; Score 25; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 WTLK 10
    ||||
DB 41 WTLK 44

```

```

RESULT 14
US-09-949-375A-12
; Sequence 12, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE

```

```
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.
US-09-949-375A-12
```

```
Query Match      69.4%; Score 25; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 WTLK 10
      ||||
Db      19 WTLK 22
```

## RESULT 15

```
US-09-864-761-41804
; Sequence 41804, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41804
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010582.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9
; OTHER INFORMATION: EST HUMAN HIT: AW873461.1, EVALUE 3.00e-18
; OTHER INFORMATION: SWISSPROT HIT: P16473, EVALUE 1.00e-27
US-09-864-761-41804
```

```
Query Match      69.4%; Score 25; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 WTLK 10
      ||||
Db      21 WTLK 24
```

Search completed: January 29, 2003, 10:59:00  
Job time : 6.54545 secs



5

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:29:22 ; Search time 29.0182 Seconds  
(without alignments)  
64.288 Million cell updates/sec

Title: US-09-707-738-24

Perfect score: 35

Sequence: 1 XXXXXXWTLKXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	71.4	4	19 AAW50120	Pan DR binding pep
2	25	71.4	8	16 AAR73328	Human TSH receptor
3	25	71.4	8	16 AAR73329	Human TSH receptor
4	25	71.4	11	21 AAY79540	Pan-DR binder pep
5	25	71.4	11	23 AAE19459	PADRE, pan-DR bind
6	25	71.4	11	23 AAU09839	Pan-DR binder pep
7	25	71.4	13	16 AAR70249	Pan DR-binding pep
8	25	71.4	13	16 AAR70250	Pan DR-binding pep
9	25	71.4	13	18 AAW22120	Padre (pan-DR bind
10	25	71.4	13	18 AAW22121	Padre (pan-DR bind

11	25	71.4	13	19 AAW50125	Pan DR binding pep
12	25	71.4	13	19 AAW50126	Pan DR binding pep
13	25	71.4	13	21 AAB36289	Promiscuous T help
14	25	71.4	13	21 AAY52558	Universal helper T
15	25	71.4	13	22 AAB73644	Pan-DR binding pep
16	25	71.4	13	22 AAB99709	Pan-DR-binding pep
17	25	71.4	13	22 AAB99710	Pan-DR-binding pep
18	25	71.4	13	22 AAB99711	Pan-DR-binding pep
19	25	71.4	13	22 AAG84519	Pan-DR-binding epi
20	25	71.4	13	22 AAG88271	Pan-DR-binding epi
21	25	71.4	13	22 AAB20154	PADRE peptide. Sy
22	25	71.4	13	22 AAB46167	PADRE universal T
23	25	71.4	13	22 AAB49066	PADRE T-cell epito
24	25	71.4	13	23 ABC01954	158P107 related HL
25	25	71.4	13	23 AAU08293	Pan DR epitope pep
26	25	71.4	13	23 AEG34860	Pan-DR-binding epi
27	25	71.4	13	23 AAU91547	Pan-DR-binding epi
28	25	71.4	13	23 ABB94471	Helper CD4 peptide
29	25	71.4	13	23 AAU10848	Padre (pan-DR bind
30	25	71.4	15	18 AAW22123	N-terminus of NADP
31	25	71.4	15	23 AAU75230	Human platelet 12-
32	25	71.4	18	20 AAW89677	Amino acids 176-19
33	25	71.4	19	16 AAR91051	SEQ ID NO 353 from
34	25	71.4	19	20 AAY19635	Tetanus toxoid epi
35	25	71.4	20	22 AAB46179	Tetanus toxoid epi
36	25	71.4	20	22 AAB46182	Human APP A-beta 1
37	25	71.4	20	22 AAB46204	Amyloid beta/PADRE
38	25	71.4	20	22 AAB49078	Amyloid beta/PADRE
39	25	71.4	20	22 AAU49081	Human cytomagalovi
40	25	71.4	22	23 AAU10836	Human cytomagalovi
41	25	71.4	22	23 AAU10839	Human cytomagalovi
42	25	71.4	22	23 AAU10842	Human cytomagalovi
43	25	71.4	22	23 AAU10843	Human cytomagalovi
44	25	71.4	22	23 AAU10834	Human cytomagalovi
45	25	71.4	24	22 ABB36700	Peptide #4206 enco

## ALIGNMENTS

RESULT 1  
AAW50120  
ID AAW50120 standard; peptide; 4 AA.  
XX AC AAW50120;  
XX 30-JUN-1998 (first entry)  
DT Pan DR binding peptide (17).  
XX Pan DR binding peptide; antigen binding site; MHC molecule;  
XX DR locus.  
OS Synthetic.  
PN US5736142-A.  
XX 07-APR-1998.  
XX 14-SEP-1994; 94US-0305871.  
XX 14-SEP-1994; 94US-0305871.  
XX 14-SEP-1993; 93US-0121101.  
XX (CYTE-) CYTEL CORP.  
PA Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
PI WPI; 1998-239154/21.  
DR Peptides that bind to MHC molecules of all DR alleles - inhibiting  
XX PT or inducing MHC Class II mediated activation of T cells

```

PS Disclosure; Columns 37-38; 29pp; English.
CC The present sequence, a pan DR binding peptide, is capable of
CC binding antigen binding sites on MHC molecules, which are encoded
CC by most of the alleles of a DR locus. The peptide can be used to
CC inhibit or induce MHC Class II mediated activation of T-cells or
CC helper T-cells, which themselves mediate a CTL response. The
CC peptide can be used in mammals, especially humans, to inhibit
CC T-cell-mediated events involved in allograft rejection, allergic
CC responses and autoimmunity and as a vaccine adjuvant for enhancing
CC an immune response against an administered immunogen. The peptide
CC can be used with other immunogens to treat, e.g. prostate cancer,
CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,
CC lymphoma, CMV and condyloma acuminatum.
XX
SQ Sequence 4 AA;
Query Match 71.4%; Score 25; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 1 WTLK 4

RESULT 2
AAR73328
ID AAR73328 standard; Peptide; 8 AA.
XX
AC AAR73328;
XX
DT 12-DEC-1995 (first entry)
XX
DE Human TSH receptor (residues 255-262).
XX
KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;
KW antibody; affinity; detection.
XX
OS Synthetic.
XX
PN JP07089991-A.
XX
PD 04-APR-1995.
XX
PF 28-SEP-1993; 93JP-0240853.
XX
PR 28-SEP-1993; 93JP-0240853.
XX
PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.
XX
DR WPI; 1995-167251/22.
XX
PT Novel polypeptide(s) having affinity for the human TSH receptor
FT antibody - used in detection of the TSH antibody.
XX
PS Claim 1; Page 23; 54pp; Japanese.
XX
CC Peptides with affinity to human TSH (thyroid stimulating hormone)
CC receptor antibody are used for detection of the antibody. (See also
CC AAR73201-592).
XX
SQ Sequence 8 AA;
Query Match 71.4%; Score 25; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 4 WTLK 7

PS Disclosure; Columns 37-38; 29pp; English.
CC The present sequence, a pan DR binding peptide, is capable of
CC binding antigen binding sites on MHC molecules, which are encoded
CC by most of the alleles of a DR locus. The peptide can be used to
CC inhibit or induce MHC Class II mediated activation of T-cells or
CC helper T-cells, which themselves mediate a CTL response. The
CC peptide can be used in mammals, especially humans, to inhibit
CC T-cell-mediated events involved in allograft rejection, allergic
CC responses and autoimmunity and as a vaccine adjuvant for enhancing
CC an immune response against an administered immunogen. The peptide
CC can be used with other immunogens to treat, e.g. prostate cancer,
CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,
CC lymphoma, CMV and condyloma acuminatum.
XX
SQ Sequence 4 AA;
Query Match 71.4%; Score 25; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 1 WTLK 4

RESULT 2
AAR73328
ID AAR73328 standard; Peptide; 8 AA.
XX
AC AAR73328;
XX
DT 12-DEC-1995 (first entry)
XX
DE Human TSH receptor (residues 255-262).
XX
KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;
KW antibody; affinity; detection.
XX
OS Synthetic.
XX
PN JP07089991-A.
XX
PD 04-APR-1995.
XX
PF 28-SEP-1993; 93JP-0240853.
XX
PR 28-SEP-1993; 93JP-0240853.
XX
PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.
XX
DR WPI; 1995-167251/22.
XX
PT Novel polypeptide(s) having affinity for the human TSH receptor
FT antibody - used in detection of the TSH antibody.
XX
PS Claim 1; Page 23; 54pp; Japanese.
XX
CC Peptides with affinity to human TSH (thyroid stimulating hormone)
CC receptor antibody are used for detection of the antibody. (See also
CC AAR73201-592).
XX
SQ Sequence 8 AA;
Query Match 71.4%; Score 25; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 4 WTLK 7

```

```

RESULT 3
AAR73329
ID AAR73329 standard; Peptide; 8 AA.
XX
AC AAR73329;
XX
DT 12-DEC-1995 (first entry)
XX
DE Human TSH receptor (residues 257-264).
XX
KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;
KW antibody; affinity; detection.
XX
OS Synthetic.
XX
PN JP07089991-A.
XX
PD 04-APR-1995.
XX
PF 28-SEP-1993; 93JP-0240853.
XX
PR 28-SEP-1993; 93JP-0240853.
XX
PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.
XX
DR WPI; 1995-167251/22.
XX
PT Novel polypeptide(s) having affinity for the human TSH receptor
FT antibody - used in detection of the TSH antibody.
XX
PS Claim 1; Page 23; 54pp; Japanese.
XX
CC Peptides with affinity to human TSH (thyroid stimulating hormone)
CC receptor antibody are used for detection of the antibody. (See also
CC AAR73201-592).
XX
SQ Sequence 8 AA;
Query Match 71.4%; Score 25; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
Db 2 WTLK 5

RESULT 4
AAY79540
ID AAY79540 standard; Peptide; 11 AA.
XX
AC AAY79540;
XX
DT 15-AUG-2000 (first entry)
XX
DE Pan-DR binder peptide PADRE 965.10.
XX
KW Antigen presenting cell; T-lymphocyte; T-cell; immunomodulator;
KW autoimmune disease; allergy; cancer; infection; graft rejection;
KW immunotherapy; therapy; pan-DR binding peptide; PADRE; antigen.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 2 /note= "cyclohexylalanine"
XX
PN WO200023053-A2.
XX
PD 27-APR-2000.
XX
PF 19-OCT-1999; 99WO-US24666.
XX

```

```

PR 20-OCT-1998; 98US-0105018.
PA (ALBA/) ALBANI S.
XX
PI Albani S;
XX
DR WPI; 2000-339492/29.
XX
PT New artificial antigen presenting cells useful for isolating and
PT expanding T cells; and modulating T cell responses for the treatment of
PT e.g. autoimmune diseases, allergies -
XX
PS Example 16; Page 93; 179pp; English.
XX
CC This synthetic peptide, termed PADRE 965.10, is characterised as a
CC pan-DR binder peptide. PADRES of comparable affinity were used as
CC model antigens to identify human antigen-specific T cells in an
CC experiment designed to demonstrate that capture of T cells by
CC artificial antigen presenting cells (APC) is effective in
CC identifying polyclonal class II restricted human T cells. The
CC invention is directed to artificial APC, and methods of making APC
CC used to isolate and expand T cell populations and to modulate
CC the identification and isolation of antigen-specific T cells. The
CC methods provide for the construction of liposomes containing
CC MHC:peptide complexes, accessory molecules, co-stimulatory
CC molecules, adhesion molecules, and other molecules irrelevant to T
CC cell binding or modulation that are used in the binding of
CC artificial APC to solid support systems that may be used in the
CC retrieval and identification and antigen-specific T cells. Devices
CC and methods are provided for treating conditions that would benefit
CC from modulation of T cell response, e.g. autoimmune disorders
CC (especially type I diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, dermatomyositis, juvenile rheumatoid
CC arthritis and uveitis), allergies, cancer, viral infections, and
CC graft rejection.
XX
SQ Sequence 11 AA;
Query Match 71.4%; Score 25; DB 21; Length 11;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB ||||
6 WTLK 9
RESULT 5
AAE19459
ID AAE19459 standard; peptide; 11 AA.
XX
AC AAE19459;
XX
DT 31-MAY-2002 (first entry)
XX
DE PADRE, pan-DR binder peptide.
XX
KW Human leukocyte antigen; HLA; DR-binding peptide; therapy;
KW stress protein; major histocompatibility complex; MHC; antiulcer;
KW type I diabetes; scleroderma; heat shock protein; hsp; vitiligo;
KW rheumatoid arthritis; lupus erythematosus; myasthenia gravis; tumour;
KW ulcerative colitis; infectious disease; haemostatic; nephrotropic;
KW polymyositis; chronic active hepatitis; primary biliary cirrhosis;
KW pernicious anaemia; autoimmune thyroiditis; Sjogren's syndrome; cancer;
KW Grave's disease; autoimmune disease; multiple sclerosis; hepatotropic;
KW inflammatory bowel disease; Goodpasture's syndrome; neuroprotective;
KW antimicrobial; immunosuppressive; dermatological; antinflammatory;
KW anticellulitis; idiopathic Addison's disease; cytostatic.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers

```

```

FT Misc-difference 2 /label= Unknown
FT WO200212286-A2.
XX
PN 14-FEB-2002.
XX
PD 08-AUG-2001; 2001WO-US41656.
XX
PF 09-AUG-2000; 2000US-224104P.
XX
PR 06-APR-2001; 2001US-0828574.
XX
PA (UYCA-) UNIV CALIFORNIA SAN DIEGO.
XX
PI Albani S, Prakkien BJ;
XX
XX WPI; 2002-227137/28.
XX
PT Novel human leukocyte antigen pan DR-binding peptide, useful for
PT treating immune mediated diseases and conditions, has a fragment of
PT stress protein that binds to major histocompatibility complex class II
PT molecules -
XX
PS Example 3; Page 46; 68pp; English.
XX
CC The invention relates to human leukocyte antigen (HLA) pan DR-binding
CC peptide comprising a fragment of a stress protein that binds to one or
CC more major histocompatibility complex (MHC) class II molecules. The
CC invention also relates to heat shock protein (hsp) peptides. The peptides
CC of the invention and thr immunomodulating composition comprising these
CC peptides are useful for modulating, treating or preventing an immune-
CC mediated disease in a mammalian subject e.g. human, having or at risk of
CC having a disease including autoimmune disease, multiple sclerosis (MS),
CC rheumatoid arthritis, lupus erythematosus, myasthenia gravis, type I
CC diabetes, scleroderma, ulcerative colitis, cancer (e.g. melanoma,
CC lymphoma, leukaemia, lung, liver, kidney, brain, bladder solid tumours,
CC retinoblastoma, sarcoma and connective tissue cancers) and infectious
CC diseases. The peptides of the invention are also useful for screening
CC peptides or analogues that modulate pathogenic immune response. These
CC peptides are useful for treating autoimmune diseases or disorders
CC including Sjogren's syndrome, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, vitiligo, idiopathic Addison's disease,
CC gluten-sensitive enteropathy, Grave's disease, inflammatory bowel
CC disease, autoimmune neutropaenia, idiopathic thrombocytopenia purpura,
CC pemphigus vulgaris, autoimmune infertility, Goodpasture's syndrome,
CC bullous pemphigoid, discoid lupus and dense deposit disease. The present
CC sequence is PADRE, a pan-DR binder peptide used in the exemplification
CC of the invention.
XX
SQ Sequence 11 AA;
Query Match 71.4%; Score 25; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 WTLK 10
DB ||||
6 WTLK 9
RESULT 6
AAU09839
ID AAU09839 standard; peptide; 11 AA.
XX
AC AAU09839;
XX
XX 26-FEB-2002 (first entry)
XX
DT Pan-DR binder peptide (PADRE).
XX
DE Antigenic; antidiabetic; neuroprotective; antirheumatic; antiarthritic;
KW dermatological; immunosuppressive; ophthalmological; antiallergic;

```

```

KW cyrostatic; virucide; antibacterial; artificial antigen presenting cell;
KW liposome; major histocompatibility complex; MHC; T cell; allograft;
KW cytokine production; T-helper 2 response; transplantation therapy;
KW autoimmune disease; type 1 diabetes mellitus; multiple sclerosis;
KW rheumatoid arthritis; dermatomyositis; juvenile rheumatoid arthritis;
KW uveitis; allergy; cancer; infection; Pan-DR binder peptide; PADRE.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /label= OTHER
FT /note= cyclohexylalanine"
XX
XX WO2001180833-A1.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2000; 2000WO-IT00161.
XX
XX 20-APR-2000; 2000WO-IT00161.
XX
XX (ALBA/) ALBANI S.
XX
XX Albani S;
XX
XX WPI; 2002-055316/07.
XX
XX New artificial antigen presenting cell, useful for modulating T cell
XX response for treating allergies and cancers, comprises liposome, major
XX histocompatibility complex, antigen and accessory molecule components -
XX
XX Example 16; Page 103; 195pp; English.
XX
XX The invention relates to an artificial antigen presenting cell (1)
XX comprising liposome, major histocompatibility complex (MHC), antigen and
XX accessory molecule components. (1) is used in methods for the
XX following: (1) identifying T cells specific for an antigen of interest;
XX (2) isolating T cells specific for an antigen of interest; (3) modulating
XX T cell response; (4) characterising the functional state of antigen-
XX specific T cells; (5) treating a condition in a subject which would be
XX benefited by altering the functional pattern of cytokine production by
XX certain antigen-specific T cells to increase T-helper (Th) 2 response
XX and/or decrease Th1 response; (6) identifying antigen-specific T cells
XX specific for epitopes on a graft donor's tissue likely to elicit graft
XX versus host rejection response; and (7) treating a recipient mammal to
XX reduce rejection of allografts in a transplantation therapy regime.
XX Method (5) is useful for treating autoimmune disease such as type I
XX diabetes mellitus, multiple sclerosis, rheumatoid arthritis,
XX dermatomyositis, juvenile rheumatoid arthritis or uveitis. Alternatively
XX it is useful for treating allergy due to allergens such as dust, animal
XX skin bypass products, vegetables, fruits, pollen or chemicals, cancer,
XX viral infection, bacterial infection. Addition of the accessory
XX molecules, as well as co-stimulatory molecules, and other proteins in
XX proper orientation in the liposomes allow for substantially improved
XX binding association and manipulation of T cells which is very important
XX in the identification and stimulation of antigen-specific T cells. The
XX present sequence represents Pan-DR binder peptide (PADRE) used in
XX the method of the invention.
XX
XX Sequence 11 AA;
XX
XX Query Match 71.4%; Score 25; DB 23; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 66;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 7 WTLK 10
XX ||||
XX Db 6 WTLK 9
XX
XX RESULT 7
XX AAR70249

```

```

ID AAR70249 standard; peptide; 13 AA.
XX
XX AAR70249;
AC
XX 13-NOV-1995 (first entry)
DT
XX Pan DR-binding peptide activator of T cells.
DE
XX T cell; MHC; class II; major histocompatibility complex II;
XX immunogen; activation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "D-form residue"
FT
XX Modified-site 3 /label= OTHER
FT /note= "cyclohexylalanine"
FT
XX Misc-difference 13 /note= "D-form residue"
FT
XX WO9507707-A.
XX
XX 23-MAR-1995.
XX
XX 14-SEP-1994; 94WO-US10368.
XX
XX 14-SEP-1993; 93US-0121101.
XX (CYTE-) CYTEL CORP.
XX
XX Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;
XX WPI; 1995-131178/17.
XX
XX Inhibiting or inducing an immune response using Pan DR-binding
XX peptide(s) - to prepare a compsn. used to treat allo:graft
XX rejection, allergic response and auto:immunity and as a vaccine
XX component
XX
XX Claim 3; Page 51; 59pp; English.
XX
XX Peptides capable of binding MHC class II molecules via antigen
XX binding sites (pan DR-binding peptides) were synthesised. Peptide
XX AAR70249 is capable of binding MHC class II molecules causing T cell
XX activation. This is useful as a vaccine component and may be used in
XX conjunction with CTL peptides to enhance a CTL response.
XX
XX Sequence 13 AA;
XX
XX Query Match 71.4%; Score 25; DB 16; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 79;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 7 WTLK 10
XX ||||
XX Db 7 WTLK 10
XX
XX RESULT 8
XX AAR70250
XX ID AAR70250 standard; peptide; 13 AA.
XX
XX AAR70250;
AC
XX 13-NOV-1995 (first entry)
DT
XX Pan DR-binding peptide activator of T cells.
DE
XX T cell; MHC; class II; major histocompatibility complex II;
XX immunogen; activation.
XX
XX

```

OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 XX WO9507707-A.  
 PN  
 XX 23-MAR-1995.  
 PD  
 XX 14-SEP-1994; 94WO-US10368.  
 PF  
 XX 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
 PI WPI; 1995-131178/17.  
 DR  
 XX Inhibiting or inducing an immune response using Pan DR-binding  
 PT peptide(s) - to prepare a compsn. used to treat allo:graft  
 PT rejection, allergic response and auto:immunity and as a vaccine  
 PT component  
 PS Claim 3; Page 51; 59pp; English.  
 XX Peptides capable of binding MHC class II molecules via antigen  
 CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
 CC AAR70250 is capable of binding MHC class II molecules causing T cell  
 CC activation. This can be used as a vaccine component. The peptide can also  
 CC be used in combination with CTL peptides to enhance a CTL response.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 71.4%; Score 25; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 Db |||||  
 7 WTLK 10

RESULT 9  
 AAW22120  
 ID AAW22120 standard; peptide; 13 AA.  
 XX  
 AC AAW22120;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 5.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Modified-site 3 /note= "cyclohexylalanine"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 XX WO9726784-A1.  
 PN  
 XX 31-JUL-1997.  
 PD  
 XX  
 XX

PF 23-JAN-1997; 97WO-US01041.  
 XX  
 DR 24-JAN-1996; 96US-0010510.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 DR  
 XX Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 FT prevention and treatment of tumours and infections  
 XX  
 PS Claim 23; Page 74; 87pp; English.  
 XX  
 CC This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 71.4%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 Db |||||  
 7 WTLK 10

RESULT 10  
 AAW22121  
 ID AAW22121 standard; peptide; 13 AA.  
 XX  
 AC AAW22121;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 6.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 XX WO9726784-A1.  
 PN  
 XX 31-JUL-1997.  
 PD  
 XX  
 XX 23-JAN-1997; 97WO-US01041.  
 PF  
 XX 24-JAN-1996; 96US-0010510.  
 PR  
 XX

PA (CYTE-) CYTEL CORP.  
 XX Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 XX Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 PT prevention and treatment of tumours and infections  
 XX  
 XX Claim 23; Page 74; 87pp; English.  
 XX This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 71.4%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 7 WTLK 10  
 RESULT 11  
 AAW50125  
 ID AAW50125 standard; peptide; 13 AA.  
 XX  
 XX AAW50125;  
 AC  
 DT 30-JUN-1998 (first entry)  
 XX  
 XX Pan DR binding peptide (22).  
 DE  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 KW  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 3 /label= Tyr, Phe  
 FT  
 XX US5736142-A.  
 PN  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.  
 XX  
 XX 14-SEP-1994; 94US-0305871.  
 XX  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 39-40; 29pp; English.  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or  
 CC helper T-cells, which themselves mediate a CTL response. The  
 CC peptide can be used in mammals, especially humans, to inhibit  
 CC T-cell-mediated events involved in allograft rejection, allergic  
 CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
 CC an immune response against an administered immunogen. The peptide  
 CC can be used with other immunogens to treat, e.g. prostate cancer,  
 CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
 CC lymphoma, CMV and condyloma acuminatum.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 71.4%; Score 25; DB 19; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 7 WTLK 10  
 RESULT 12  
 AAW50126  
 ID AAW50126 standard; peptide; 13 AA.  
 XX  
 XX AAW50126;  
 AC  
 DT 30-JUN-1998 (first entry)  
 XX  
 XX Pan DR binding peptide (23).  
 DE  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 KW  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 3 /label= Tyr, Phe  
 FT  
 XX US5736142-A.  
 PN  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.  
 XX  
 XX 14-SEP-1994; 94US-0305871.  
 XX  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 PR  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.  
 XX  
 XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 41-42; 29pp; English.  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or

CC helper T-cells, which themselves mediate a CTL response. The  
CC peptide can be used in mammals, especially humans, to inhibit  
CC T-cell-mediated events involved in allograft rejection, allergic  
CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
CC an immune response against an administered immunogen. The peptide  
CC can be used with other immunogens to treat, e.g. prostate cancer,  
CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
CC lymphoma, CMV and condyloma acuminatum.  
XX  
SQ Sequence 13 AA;

Query Match 71.4%; Score 25; DB 19; Length 13;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
DB |||||  
7 WTLK 10

RESULT 13  
AAB36289  
ID AAB36289 standard; Peptide; 13 AA.  
XX AC AAB36289;  
XX  
XX  
DT 26-FEB-2001 (first entry)  
XX  
DE Promiscuous T helper epitope SEQ ID NO: 65.  
XX  
XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;  
KW cancer; eosinophilia; vaccine; allergic rhinitis.  
XX  
XX Synthetic.  
OS  
XX WO200065058-A1.  
PN  
XX 02-NOV-2000.  
PD  
XX 19-APR-2000; 2000WO-DK00205.  
PF  
XX 23-APR-1999; 99DK-0000552.  
PR 06-MAY-1999; 99US-0132811.  
XX  
XX (MBBI-) M & E BIOTECH AS.  
PA  
XX Klysner S;  
PI  
XX WPI; 2000-672791/65.  
DR  
XX Down-regulating interleukin 5 (IL-5) activity in humans by  
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,  
PT prophylaxis or amelioration of asthma or other chronic allergic  
PT conditions -  
XX  
XX Disclosure; Page 168; 172pp; English.  
PS  
XX The present invention is concerned with methods of treating asthma,  
CC eosinophilia, allergic rhinitis and other allergic diseases. These  
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
CC proteins and their coding sequences to down-regulate IL-5 activity and  
CC thus reduce eosinophil numbers. The allergic diseases may be treated  
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
CC it is possible that they may be used in the treatment of cancer and  
CC helminthic infections.  
XX  
XX Sequence 13 AA;

Query Match 71.4%; Score 25; DB 21; Length 13;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db |||||  
7 WTLK 10

RESULT 14  
AAY52558  
ID AAY52558 standard; peptide; 13 AA.  
XX AC AAY52558;  
XX  
XX  
DT 28-FEB-2000 (first entry)  
XX  
DE Universal helper T epitope, pan DR epitope (PADRE).  
XX  
KW Chimeric; II protein; pan DR epitope; expression vector;  
KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
KW epitope; antigen; presentation; class I; cytosolic pathway;  
KW endoplasmic reticulum; class II; extracellular antigen;  
KW endocytic pathway; helper T lymphocyte; HII; universal epitope;  
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
KW autoimmune disease; activation; antiviral; antimalarial;  
KW immunoprotective.  
XX  
XX Synthetic.  
OS  
XX WO9958658-A2.  
PN  
XX 18-NOV-1999.  
PD  
XX 13-MAY-1999; 99WO-US10646.  
PF  
XX 13-MAY-1998; 98US-0078904.  
PR 15-MAY-1998; 98US-0085751.  
XX  
XX (EPIM-) EPIMMUNE INC.  
PA  
XX Fikes JD, Hermanson GG, Sette A, Ishioka GV, Livingston B;  
PI Chesnut RW;  
PI  
XX WPI; 2000-039103/03.  
DR N-PSDB; AAZ38679.  
XX  
XX Expression vectors encoding major histocompatibility targeting  
PT sequence, used as, e.g. tumor vaccines -  
XX  
XX Claim 9; Page 80; 130pp; English.  
PS  
XX This sequence represents a universal helper T epitope, pan DR epitope  
CC (PADRE), DNA encoding which is used to construct fusion genes used  
CC in exemplifications of the present invention. The invention  
CC relates to a novel expression vector comprising a promoter operably  
CC linked to a fusion gene encoding a major histocompatibility complex  
CC (MHC) targeting sequence, and two or more heterologous peptide epitopes.  
CC The MHC targeting sequence may be a class I targeting sequence, which  
CC directs an MHC class I epitope to a cytosolic pathway or to the  
CC endoplasmic reticulum, or an MHC class II targeting sequence, which  
CC directs extracellular antigens to enter the endocytic pathway to be  
CC processed into antigen peptides for presentation on MHC class II  
CC molecules. The heterologous epitopes may comprise either T  
CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and  
CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are  
CC useful for stimulating an immune response in vivo, as well as for use in  
CC assaying the human immunogenicity of a human T cell peptide epitope in  
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
CC combined with other MHC class I and class II epitopes to increase the  
CC number of cells that are activated in response to a given antigen and  
CC provide a broader population coverage of MHC-reactive alleles.

XX SQ Sequence 13 AA;  
 CC Query Match 71.4%; Score 25; DB 21; Length 13;  
 CC Best Local Similarity 100.0%; Pred. No. 79;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX QY 7 WTLK 10  
 DB 7 WTLK 10  
 RESULT 15  
 AAB73644  
 ID AAB73644 standard; peptide: 13 AA.  
 XX AC AAB73644;  
 XX DT 11-SEP-2001 (first entry)  
 XX DE Pan-DR binding peptide, PADRE.  
 XX KW Cellular vaccine; antigen-presenting cell; APC;  
 KW cell-surface molecule density; major histocompatibility complex; MHC;  
 KW antigen-specific T-cell expansion; tumour; cancer; viral infection;  
 KW parasitic infection; Pan-DR binding peptide; PADRE.  
 XX OS Unidentified.  
 XX WO200136978-A1.  
 XX PN 25-MAY-2001.  
 XX PD 15-NOV-2000; 2000WO-US42213.  
 XX PF 15-NOV-1999; 99US-0165428.  
 XX PR (DAND ) DANA FARBER CANCER INST INC.  
 XX PA Schultze JL, Vonderheide RH, Nadler LM, Maecker B;  
 XX PI Von Bergwelt-Baildon M;  
 XX DR WPI; 2001-343909/36.  
 XX PT Determining surface density of molecules on antigen-presenting cells,  
 PT useful for quality control of cellular vaccines for treatment of tumors  
 PT and infections comprises measuring density ratios -  
 XX PS Disclosure; Page 21; 47pp; English.  
 XX CC The invention relates to a method of determining the surface density of  
 CC a cell-surface molecule on a primary or artificial antigen-presenting  
 CC cell (APC). The method comprises determining the cell surface area  
 CC of the APCs, determining the absolute amount of cell surface molecule on  
 CC the surface of the APCs, and calculating the ratio of the amount of  
 CC the cell surface molecule to the APC surface area as a measure of cell  
 CC surface molecule density. The invention also encompasses a kit for  
 CC determining the cell-surface density of a complex of peptide and MHC  
 CC (major histocompatibility complex) protein/peptide complex on an APC;  
 CC identifying an agent that increases persistence of the MHC/peptide  
 CC complex on an APC surface by culturing APCs in presence of test compound  
 CC and measuring the time of persistence relative to an untreated control;  
 CC and determining if the APC carries a therapeutically adequate amount of  
 CC peptide by determining whether the cell-surface density of the  
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The  
 CC method is used to determine if the amount of MHC/antigenic peptide  
 CC complex present on the surface of an APC is sufficient for therapeutic  
 CC use of the cells as cellular vaccines or for ex vivo expansion of  
 CC antigen-specific T cells for subsequent return to the patient,  
 CC particularly for the treatment of tumours or viral or parasitic  
 CC infections. The method can also be used to identify agents that increase  
 CC production of MHC/peptide complexes on cells, (which increases the  
 CC therapeutic potential of the cells) by treatment before and/or during

CC APC administration. The method allows identification of therapeutically  
 CC useful antigen-pulsed APC optimisation of conditions for their  
 CC production, and control of APC quality. Sequences AAB73643-AAB73648  
 CC represent peptides used in MHC binding studies and for the generation  
 CC of peptide-specific cytotoxic T-lymphocytes.

XX SQ Sequence 13 AA;

Query Match 71.4%; Score 25; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 DB 7 WTLK 10

Search completed: January 29, 2003, 10:44:34  
 Job time : 29.0182 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:46:52 ; Search time 6.10909 Seconds  
(without alignments)  
46.243 Million cell updates/sec

Title: US-09-707-738-24

Perfect score: 35

Sequence: 1 XXXXXWTLKXXXX 14

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:

1: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW PUB.pap.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW PUB.pap.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pap.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pap.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pap.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pap.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pap.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	71.4	11	10	US-09-756-983-7
2	25	71.4	11	10	US-09-828-574-12
3	25	71.4	13	9	US-09-949-375A-18
4	25	71.4	13	9	US-09-785-215-19
5	25	71.4	13	10	US-09-894-018-69
6	25	71.4	24	10	US-09-864-761-47626
7	25	71.4	24	10	US-09-983-019-10
8	25	71.4	27	10	US-09-864-761-40242
9	25	71.4	27	10	US-09-983-019-4
10	25	71.4	27	10	US-09-983-019-7
11	25	71.4	28	10	US-09-864-761-39731
12	25	71.4	48	10	US-09-864-761-45303
13	25	71.4	51	9	US-10-016-634A-165
14	25	71.4	55	9	US-09-949-375A-12
15	25	71.4	56	10	US-09-864-761-41804
16	25	71.4	57	10	US-09-864-761-44535
17	25	71.4	58	10	US-09-925-300-1859
18	25	71.4	59	10	US-09-925-299-822
19	25	71.4	62	10	US-09-896-578-5

20	25	71.4	73	10	US-09-864-761-47212	Sequence 47212, A
21	25	71.4	74	10	US-09-764-877-1943	Sequence 1943, Ap
22	25	71.4	80	10	US-09-894-018-97	Sequence 97, Appl
23	25	71.4	87	10	US-09-925-297-562	Sequence 562, App
24	25	71.4	89	10	US-09-867-550-686	Sequence 686, App
25	25	71.4	97	10	US-09-864-761-43929	Sequence 43929, A
26	25	71.4	98	10	US-09-894-018-101	Sequence 101, App
27	25	71.4	106	9	US-10-138-536-3	Sequence 3, Appl
28	25	71.4	106	9	US-10-146-130-5	Sequence 5, Appl
29	25	71.4	106	10	US-09-894-018-95	Sequence 95, Appl
30	25	71.4	107	10	US-09-894-018-93	Sequence 93, Appl
31	25	71.4	107	10	US-09-894-018-103	Sequence 103, App
32	25	71.4	109	10	US-09-910-150-32	Sequence 32, Appl
33	25	71.4	113	9	US-09-789-054A-18	Sequence 18, Appl
34	25	71.4	123	10	US-09-894-018-109	Sequence 109, App
35	25	71.4	130	10	US-09-894-018-99	Sequence 99, Appl
36	25	71.4	144	10	US-09-894-018-129	Sequence 129, App
37	25	71.4	147	10	US-09-894-018-131	Sequence 131, App
38	25	71.4	148	10	US-09-894-018-127	Sequence 127, App
39	25	71.4	152	10	US-09-939-980-264	Sequence 264, App
40	25	71.4	157	10	US-09-894-018-117	Sequence 117, App
41	25	71.4	168	10	US-09-894-018-115	Sequence 115, App
42	25	71.4	169	10	US-09-894-018-119	Sequence 119, App
43	25	71.4	170	9	US-09-479-040-3	Sequence 3, Appl
44	25	71.4	180	10	US-09-894-018-125	Sequence 125, App
45	25	71.4	183	9	US-09-883-152-19	Sequence 19, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-756-983-7  
; Sequence 7, Application US/09756983  
; Patent No. US20020122818A1  
; GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthesized peptide totally artificial  
; OTHER INFORMATION: Xaa in position 2 stands for cyclohexylalanine  
US-09-756-983-7

Query Match 71.4%; Score 25; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 6 WTLK 9

#### RESULT 2

US-09-828-574-12  
; Sequence 12, Application US/09828574  
; Patent No. US20020146759A1

```
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: ALBANI, Salvatore
; APPLICANT: PRAKSEN, Berent J.
; TITLE OF INVENTION: STRESS PROTEINS AND PEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: UCSD1110-1
; CURRENT APPLICATION NUMBER: US/09/828,574
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/224,104
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: pan-DR binder peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is any amino acid
US-09-828-574-12

Query Match          71.4%; Score 25; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 6 WTLK 9

RESULT 3
US-09-949-375A-18
; Sequence 18, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.
US-09-949-375A-18

Query Match          71.4%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 4
US-09-785-215-19
; Sequence 19, Application US/09785215
; Publication No. US2002019157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: of MHC Class II molecules in a variety of animals
US-09-785-215-19

Query Match          71.4%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 5
US-09-894-018-69
; Sequence 69, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Deniss
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69

Query Match          71.4%; Score 25; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 6
US-09-864-761-47626
; Sequence 47626, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Weisheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; CURRENT APPLICATION NUMBER: US/09/864,761
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
```

```
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47626
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011502.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: AI033706.1, EVALUATE 1.00e-08
; OTHER INFORMATION: SWISSPROT HIT: Q84291, EVALUATE 4.90e+00
US-09-864-761-47626

Query Match 71.4%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 7 WTLK 10
Db 13 WTLK 16

RESULT 7
US-09-983-019-10
; Sequence 10, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP1E
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (3)..(3)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-10

Query Match 71.4%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10
Db 7 WTLK 10

RESULT 8
US-09-864-761-40242
; Sequence 40242, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Weisheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40242
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC025746.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-09-864-761-40242

```

```

Query Match          71.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 7 WTLK 10
    ||||
Db 1 WTLK 4

```

## RESULT 9

```
US-09-983-019-4
```

```

; Sequence 4, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (6) .. (6)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-4

```

```

Query Match          71.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 7 WTLK 10
    ||||
Db 10 WTLK 13

```

## RESULT 10

```
US-09-983-019-7
```

```

; Sequence 7, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US

```

```

; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (17)..(17)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-7

```

```

Query Match          71.4%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 7 WTLK 10
    ||||
Db 21 WTLK 24

```

## RESULT 11

```
US-09-864-761-39731
```

```

; Sequence 39731, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Bank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687

```

; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 39731  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC003681.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8  
US-09-864-761-39731

Query Match 71.4%; Score 25; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 11 WTLK 14

## RESULT 12

US-09-864-761-45303  
; Sequence 45303, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmics-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45303  
; LENGTH: 48  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC022216.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.71  
; OTHER INFORMATION: SWISSPROT HIT: Q9ZER9, EVALUATION 1.20e+00  
US-09-864-761-45303

Query Match 71.4%; Score 25; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 2 WTLK 5

## RESULT 13

US-10-016-634A-165  
; Sequence 165, Application US/10016634A  
; Publication No. US20020192666A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Ghosh, Malavika  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Prot  
; FILE REFERENCE: DEX-0255  
; CURRENT APPLICATION NUMBER: US/10/016.634A  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US 60/244,258  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 176  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 165  
; LENGTH: 51  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-016-634A-165

Query Match 71.4%; Score 25; DB 9; Length 51;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 41 WTLK 44

RESULT 14  
US-09-949-375A-12  
; Sequence 12, Application US/09949375A  
; Patent No. US20020172673A1  
; GENERAL INFORMATION:  
; APPLICANT: KLYSNER, Steen et al.  
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE

```
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 12
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.
US-09-949-375A-12
```

```
Query Match 71.4%; Score 25; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 WTLK 10
|||
Db 19 WTLK 22
```

```
RESULT 15
US-09-864-761-41804
; Sequence 41804, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41804
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010582.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9
; OTHER INFORMATION: EST HUMAN HIT: AW673461.1, EVALUE 3.00e-18
; OTHER INFORMATION: SWISSPROT HIT: P16473, EVALUE 1.00e-27
US-09-864-761-41804
```

```
Query Match 71.4%; Score 25; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 WTLK 10
|||
Db 21 WTLK 24
```

```
Search completed: January 29, 2003, 10:59:00
Job time : 6.10909 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:07 ; Search time 10.8727 Seconds  
(without alignments)  
114.943 Million cell updates/sec

Title: US-09-707-738-23  
Perfect score: 34  
Sequence: 1 XXXXXWTLKXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	25	73.5	14	2 A47421	leukotriene B-4 12
2	25	73.5	70	2 AB1984	hypothetical prote
3	25	73.5	81	2 S35586	high potential iro
4	25	73.5	85	1 IHKREV	high potential iro
5	25	73.5	85	1 IHTFER	high potential iro
6	25	73.5	90	2 D97355	hypothetical prote
7	25	73.5	96	2 S23863	hypothetical prote
8	25	73.5	104	2 A41384	nicotinic acetylch
9	25	73.5	122	2 A45385	translation repres
10	25	73.5	127	2 C97299	hypothetical prote
11	25	73.5	134	2 T12858	hypothetical prote
12	25	73.5	137	2 G90472	partial transposas
13	25	73.5	152	2 F90143	hypothetical prote
14	25	73.5	152	2 C90295	hypothetical prote
15	25	73.5	159	2 E72731	hypothetical prote
16	25	73.5	160	2 T93932	hypothetical prote
17	25	73.5	163	2 S69658	conserved hypotet
18	25	73.5	164	2 H89784	hypothetical prote
19	25	73.5	172	2 G89764	conserved hypotet
20	25	73.5	173	2 A83965	hypothetical prote
21	25	73.5	175	2 T48641	transposase (21) B
22	25	73.5	177	2 T00789	hypothetical prote
23	25	73.5	180	2 T46695	ubiquitin-protein
24	25	73.5	187	2 E69170	hypothetical prote
25	25	73.5	201	2 F97063	hypothetical prote
26	25	73.5	201	2 AD3302	surfactin biosynth
27	25	73.5	201	2 C97558	31K outer-membrane
28	25	73.5	201	2 AG2778	hypothetical prote
29	25	73.5	203	2 C69063	conserved hypotet

#### ALIGNMENTS

##### RESULT 1

A47421  
leukotriene B-4 12-hydroxydehydrogenase (EC 1.1.1.1.-) - pig (fragment)  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 19-May-1995 #sequence\_revision 19-May-1995 #text\_change 17-Mar-1999  
C/Accession: A47421  
R/Yokomizo, T.; Izumi, T.; Takahashi, T.; Kasama, T.; Kobayashi, Y.; Sato, F.; Taketani, J. Biol. Chem. 268, 18128-18135, 1993  
A/Title: Enzymatic inactivation of leukotriene B-4 by a novel enzyme found in the porcine  
A/Reference number: A47421; MUID:93352633; PMID:8394361  
A/Accession: A47421  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-14 <YOK>  
C/Keywords: oxidoreductase

Query Match 73.5%; Score 25; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 6 WTLK 9  
|||||

##### RESULT 2

AB1984  
hypothetical protein asl1421 [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp.  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C/Accession: AB1984  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AB1984  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-70 <KUR>  
A/Cross-references: GB:BA000019; PIDN:BAB73378.1; PID:gl7130768; GSPDB:GN00179  
A/Experimental source: strain PCC 7120  
C/Genetics:  
A/Gene: asl1421

Query Match 73.5%; Score 25; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||||

hypothetical prote  
SSU ribosomal prot  
E4 protein - cotto  
hypothetical prote  
hypothetical prote  
31K outer-membrane  
proteasome activat  
proteasome activat  
hypothetical prote  
hypothetical prote  
hypothetical prote  
ricin E - castor b  
conserved hypotet  
hypothetical prote  
probable permease  
hypothetical prote

30 25 73.5 207 2 C70029  
31 25 73.5 211 2 H69539  
32 25 73.5 212 1 W4WLRB  
33 25 73.5 217 2 AF1898  
34 25 73.5 218 2 D71693  
35 25 73.5 219 2 AC3615  
36 25 73.5 238 2 167638  
37 25 73.5 239 2 153518  
38 25 73.5 240 2 T45814  
39 25 73.5 249 2 T23077  
40 25 73.5 252 2 D87397  
41 25 73.5 263 2 S06330  
42 25 73.5 264 2 B95259  
43 25 73.5 264 2 E98124  
44 25 73.5 266 2 H98208  
45 25 73.5 266 2 AI3077

Db 17 WTLK 20

RESULT 3  
S35586  
high potential iron-sulfur protein - Chromatium tepidum  
C;Species: Chromatium tepidum  
C;Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 18-Jun-1999  
A;Accession: S35586  
R;Moullis, J.M.; Scherrer, N.; Gagnon, J.; Forest, E.; Petillot, Y.; Garcia, D.  
Arch. Biochem. Biophys. 305, 186-192, 1993  
A;Title: Primary structure of Chromatium tepidum high-potential iron-sulfur protein in  
A;Reference number: S35586; MUID:93343628; PMID:8393645  
A;Accession: S35586  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-81 <MOU>  
C;Superfamily: high potential iron-sulfur protein  
C;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein  
F;43.46.61.75/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 73.5%; Score 25; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 78 WTLK 81

RESULT 4  
IHKREV  
high potential iron-sulfur protein [validated] - Chromatium vinosum  
N;Alternate names: HiPiP  
C;Species: Chromatium vinosum  
C;Date: 24-Apr-1984 #sequence\_revision 24-Oct-1997 #text\_change 15-Sep-2000  
C;Accession: A92330; A92143; A00263  
R;Tetro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.  
J. Biol. Chem. 256, 731-735, 1981  
A;Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur  
A;Reference number: A92330; MUID:81094036; PMID:7451471  
A;Accession: A92330  
A;Molecule type: protein  
A;Residues: 1-73, 'D', '75-85 <TED>  
R;Dus, K.; Tetro, S.; Bartsch, R.G.  
J. Biol. Chem. 248, 7318-7331, 1973  
A;Title: The complete amino acid sequence of Chromatium high potential iron sulfur prote  
A;Reference number: A92143; MUID:74012043; PMID:4745771  
A;Accession: A92143  
A;Molecule type: protein  
A;Residues: 1-10, 'N', '12-44, 'D', '46-85 <DUS>  
A;Experimental source: strain D  
R;Banci, L.; Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Sompornpisut, P.  
submitted to the Brookhaven Protein Data Bank, January 1995  
A;Reference number: A65814; PDB:1HRO  
R;Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Sompornpisut, P.  
submitted to the Brookhaven Protein Data Bank, December 1995  
A;Reference number: A66207; PDB:1NEH  
A;Contents: annotation; conformation by (1)H-NMR, oxidized form, residues 1-85  
R;Bacques, G.; Mino, Y.; Loehr, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman  
J. Am. Chem. Soc. 113, 2055-2064, 1991  
A;Title: The environment of Fe4S4 clusters in ferredoxins and high-potential iron prote  
A;Reference number: A44688  
A;Contents: annotation; X-ray crystallography, 2.0 angstroms  
A;Notes: assignment of Raman spectra frequencies and hydrogen bonds around the iron-sulfu  
R;Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Xuong, N.H.; Alden, R.A.; Bartsch, R.G.  
J. Biol. Chem. 249, 4212-4225, 1974  
A;Title: Two-angstrom crystal structure of oxidized Chromatium high potential iron prote  
A;Reference number: A92153; MUID:74309824; PMID:4855287  
A;Contents: annotation; X-ray crystallography, 2.0 angstroms  
R;Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Alden, R.A.  
J. Biol. Chem. 249, 6339-6346, 1974

A;Title: Comparison of oxidation-reduction site geometries in oxidized and reduced Chroma  
A;Reference number: A92161; MUID:75019502; PMID:4417854

A;Contents: annotation; X-ray crystallography  
A;Note: structures of the oxidized and reduced forms are compared with each other and wit  
C;Superfamily: high potential iron-sulfur protein  
C;Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein  
F;43.46.63.77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental

Query Match 73.5%; Score 25; DB 1; Length 85;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 80 WTLK 83

## RESULT 5

IHTFER  
high potential iron-sulfur protein - Thiocapsa roseopersicina (tentative sequence)

C;Species: Thiocapsa roseopersicina  
C;Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 31-Mar-2000  
C;Accession: A00264  
R;Tetro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.  
J. Biol. Chem. 256, 731-735, 1981

A;Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur  
A;Reference number: A92330; MUID:81094036; PMID:7451471  
A;Accession: A00264  
A;Molecule type: protein  
A;Residues: 1-85 <TED>

C;Superfamily: high potential iron-sulfur protein; metalloprotein  
C;Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein  
F;43.46.63.77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 73.5%; Score 25; DB 1; Length 85;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 80 WTLK 83

## RESULT 6

D97355  
hypothetical protein CAC3711 [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C;Accession: D97355  
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: D97355  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-90 <KUR>

A;Cross-references: CB:AE001437; PIDN:AAK81631.1; PID:G15026816; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC3711

Query Match 73.5%; Score 25; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
Db 46 WTLK 49

## RESULT 7



## S23863

hypothetical protein 2 - garlic mosaic virus  
 C:Species: Garlic mosaic virus, GarMV  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
 C:Accession: S23863  
 R:Choi, J.N.; Choi, Y.H.; Choi, Y.D.; Lee, J.S.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Nucleotide sequence of a cDNA for garlic mosaic virus.  
 A:Reference number: S23862  
 A:Accession: S23863  
 A:Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 1-96 <CHO>  
 A:Cross-references: EMBL:X67134; NID:g59391; PIDN:CAA47614.1; PID:g59393

Query Match 73.5%; Score 25; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

## RESULT 8

A41384  
 nicotinic acetylcholine receptor alpha chain - Chinese cobra (fragment)  
 C:Species: Naja naja atra (Chinese cobra)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: A41384  
 R:Neumann, D.; Barchan, D.; Horowitz, M.; Kochva, E.; Fuchs, S.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7255-7259, 1989  
 A:Title: Snake acetylcholine receptor: cloning of the domain containing the four extracellular loops  
 A:Reference number: A41384; MUID:89386720; PMID:2780569  
 A:Accession: A41384  
 A:Molecule type: mRNA  
 A:Residues: 1-104 <NEU>  
 A:Cross-references: GB:M26388; NID:g213370; PIDN:AAA49384.1; PID:g213371  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: neurotransmitter receptor

Query Match 73.5%; Score 25; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 9

A45385  
 translation repressor regA - phage RB69  
 C:Species: Phage RB69  
 C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: A45385  
 R:Jozwik, C.E.; Miller, E.S.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 5053-5057, 1992  
 A:Title: Regions of bacteriophage T4 and RB69 RegA translational repressor proteins that interact with the 3' ends of the 16S and 23S ribosomal subunits  
 A:Reference number: A45385; MUID:92279264; PMID:1594613  
 A:Accession: A45385  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-122 <JOZ>  
 A:Cross-references: GB:M86231; NID:g215354; PIDN:AAA32295.1; PID:g215356  
 C:Genetics:  
 A:Gene: regA  
 C:Superfamily: phage T4 translation repressor  
 C:Keywords: RNA binding; translation repressor

Query Match 73.5%; Score 25; DB 2; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 112 WTLK 115

## RESULT 10

C97299  
 hypothetical protein CAC3248 [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: C97299  
 R:Nolling, J.; Breston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Dally, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: C97299  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-127 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK81182.1; PID:g15026321; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3248

Query Match 73.5%; Score 25; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 103 WTLK 106

## RESULT 11

T12858  
 hypothetical protein yopX - Bacillus subtilis phage SPBc2  
 C:Species: Bacillus subtilis phage SPBc2  
 C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: T12858; F69918  
 R:Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Muel, C.; Karamata, D.  
 submitted to the EMBL Data Library, August 1997  
 A:Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 prophage  
 A:Reference number: Z17583  
 A:Accession: T12858  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-134 <LAZ>  
 A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025572; PIDN:AAC13067.1  
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallucci, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Iech, J.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, K.; Lauher, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A9580; MUID:98044033; PMID:9384377  
 A:Accession: F69918  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-134 <KUN>  
 A:Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB13991.1; PII  
 A:Experimental source: strain 168  
 C:Genetics:

A:Gene: yopX  
C:Superfamily: Bacillus subtilis phage sPBc2 hypothetical protein yopX

Query Match 73.5%; Score 25; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 31 WTLK 34

## RESULT 12

G90472  
partial transposase ISC1078 [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: G90472  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: G90472  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-137 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3816304; PIDN:AAK43038.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO2933

Query Match 73.5%; Score 25; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

## RESULT 13

F90473  
hypothetical protein SSO0040 [imported] - Sulfolobus solfataricus transposon ISC1078  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 18-Jul-2001  
C:Accession: F90143; C90336; C90480; G90280  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: F90143  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3813171; PIDN:AAK40405.1; GSPDB:GN00155  
C:Accession: C90336  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KU2>  
A:Cross-references: GB:AE006641; NID:gl3814994; PIDN:AAK41946.1; GSPDB:GN00155  
C:Accession: C90480  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KU3>  
A:Cross-references: GB:AE006641; NID:gl3816382; PIDN:AAK43098.1; GSPDB:GN00155  
C:Accession: G90280  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KU4>  
A:Cross-references: GB:AE006641; NID:gl3814460; PIDN:AAK41502.1; GSPDB:GN00155

C:Genetics:  
A:Gene: SSO0040; SSO1750; SSO2992; SSO1264

Query Match 73.5%; Score 25; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

## RESULT 14

C90295  
hypothetical protein SSO1384 [imported] - Sulfolobus solfataricus transposon ISC1078  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: C90295  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: C90295  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3814600; PIDN:AAK41618.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO1384

Query Match 73.5%; Score 25; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

## RESULT 15

E72731  
hypothetical protein APE0390 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: E72731  
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha-  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyru  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: E72731  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-159 <KAW>  
A:Cross-references: DDBJ:AP000059; NID:g5103911; PIDN:BAA79345.1; PID:g5104029  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0390

Query Match 73.5%; Score 25; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 37 WTLK 40

Search completed: January 29, 2003, 10:47:36  
Job time : 12.8727 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:30:17 ; Search time 5,90909 Seconds  
(without alignments)  
91.248 Million cell updates/sec

Title: US-09-707-738-23  
Perfect score: 34  
Sequence: 1 XXXXXXWTLKXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	73.5	83	1 HPIS_THETI	P80176 thermochrom
2	25	73.5	85	1 HPIS_THIRO	P00261 thiocapae r
3	25	73.5	90	1 Y1B1_CLOAB	Q04353 clostridium
4	25	73.5	98	1 NULM_BRARE	Q9miy2 brachydanio
5	25	73.5	104	1 ACHA_NAJNA	P14143 najja najja (
6	25	73.5	114	1 ACPS_MYCGE	Q9zb79 mycoplasma
7	25	73.5	122	1 HPIS_CHRVI	P00260 chromatium
8	25	73.5	122	1 REGA_BPR69	Q01751 bacterioph
9	25	73.5	125	1 CP03_HUMAN	O95177 homo sapien
10	25	73.5	211	1 RS3A_ARCFU	O27964 archaeoglob
11	25	73.5	212	1 V84_CRPVK	P03124 cottontail
12	25	73.5	218	1 Y363_RICPR	Q9zdg6 rickettsia
13	25	73.5	233	1 MGRA_BOVIN	P54149 bos taurus
14	25	73.5	238	1 PSE2_RAT	Q63798 rattus norv
15	25	73.5	239	1 PSE2_HUMAN	Q9ul46 homo sapien
16	25	73.5	239	1 PSE2_MOUSE	P97372 mus musculu
17	25	73.5	240	1 OM31_BRUME	Q45322 brucellame
18	25	73.5	282	1 YC80_GUITH	O78449 guillardia
19	25	73.5	294	1 UP08_ECOLI	P39173 escherichia
20	25	73.5	301	1 LEC1_RAT	P08290 rattus norv
21	25	73.5	309	1 YC30_ODOSI	P49518 odontella s
22	25	73.5	311	1 LB4D_HUMAN	Q14914 homo sapien
23	25	73.5	329	1 LB4D_PIG	Q29073 sus scrofa
24	25	73.5	334	1 YEDY_ECO57	Q6xb74 escherichia
25	25	73.5	334	1 YEDY_ECOLI	P76342 escherichia
26	25	73.5	334	1 YEDY_SALTY	Q6xes1 salmonella
27	25	73.5	345	1 ARGC_BACHD	Q9k8v2 bacillus ha
28	25	73.5	345	1 CATV_GVXN	Q9pyy5 xestia c-ni
29	25	73.5	349	1 I10S_MOUSE	Q61190 mus musculu
30	25	73.5	349	1 LB4D_RABIT	Q28719 oryctolagus
31	25	73.5	349	1 MRAY_CHLPN	Q9z706 chlamydia p
32	25	73.5	349	1 Y028_BORBU	O51059 borrelia bu
33	25	73.5	352	1 IIBC_HUMAN	P55103 homo sapien

34	25	73.5	353	1 PSBD_MAIZE	P48184 zea mays (m
35	25	73.5	355	1 VAL1_ABMVW	P21947 abutilon mo
36	25	73.5	357	1 O2B2_HUMAN	Q9gzk3 homo sapien
37	25	73.5	360	1 VE2_HPV70	P50773 human papil
38	25	73.5	361	1 VAL1_PYMVW	P27258 potato yell
39	25	73.5	365	1 HHP1_SCHPO	P40235 schizosacch
40	25	73.5	370	1 VE2_HPV39	P24830 human papil
41	25	73.5	372	1 3BHS_MACVU	P27365 m 3 beta-hy
42	25	73.5	372	1 AR1B_HUMAN	O15143 homo sapien
43	25	73.5	372	1 AR1B_MOUSE	Q9qv32 mus musculu
44	25	73.5	372	1 AR1B_RAT	O8656 rattus norv
45	25	73.5	373	1 NOLL_RHILO	Q52778 rhizobium 1

#### ALIGNMENTS

RESULT 1  
HPIS\_THETI  
ID HPIS\_THETI STANDARD; PRT; 83 AA.  
AC P80176;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE High potential iron-sulfur protein (HiPIP).  
GN HIP.  
OS Thermochromatium tepidum (Chromatium tepidum).  
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
OC Thermochromatium.  
OX NCBI\_TaxID=1050;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=ATCC 43061;  
RX MEDLINE=93343628; PubMed=8393645;  
RA Moulis J.-M., Scherrer N., Gagnon J., Forest E., Petillot Y.,  
RA Garcia D.;  
RT "Primary structure of Chromatium tepidum high-potential iron-sulfur  
protein in relation to thermal denaturation.";  
RL Arch. Biochem. Biophys. 305:186-192(1993).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).  
RX MEDLINE=20558546; PubMed=1095707;  
RA Nogi T., Fatthir I., Kobayashi M., Nozawa T., Miki K.;  
RT "Crystal structures of photosynthetic reaction center and high-  
potential iron-sulfur protein from Thermochromatium tepidum:  
thermostability and electron transfer.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:13561-13566(2000).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (0.8 ANGSTROMS).  
RX MEDLINE=22071404; PubMed=12077426;  
RA Liu L., Nogi T., Kobayashi M., Nozawa T., Miki K.;  
RT "Ultra-high-resolution structure of high-potential iron-sulfur  
protein from Thermochromatium tepidum.";  
RL Acta Crystallogr. D 58:1085-1091(2002).  
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
FERREDXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
ONE GENUS (PARACOCUS) OF HALOPHILIC, DENTRIFYING BACTERIA.  
CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
CC -!- MISCELLANEOUS: THIS PROTEIN IS THERMOSTABLE.  
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
(HIPIP) FAMILY.  
DR PIR; S35586; S35586.  
DR PDB; 1BYT; 13-DEC-00.  
DR PDB; 1IUA; 20-MAR-02.  
DR InterPro; IPR000170; Hipot\_ironsul.  
DR Pfam; PF01355; HiPIP; 1.  
DR PRINTS; PR00374; HIPIPRDXIN.  
DR PROSITE; PS00596; HiPIP; 1.  
KW Electron transport; Iron-sulfur; 4Fe-4S; 3D-structure.  
FT METAL 43 43 IRON-SULFUR (4FE-4S).  
FT METAL 46 46 IRON-SULFUR (4FE-4S).

FT METAL 61 61 IRON-SULFUR (4FE-4S).  
 FT METAL 75 75 IRON-SULFUR (4FE-4S).  
 SQ SEQUENCE 83 AA; 8786 MW; 9211684FD2C44E0A CRC64;

Query Match 73.5%; Score 25; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 DB 78 WTLK 81

## RESULT 2

ID HPIS\_THIRO STANDARD; PRT; 85 AA.

AC P00261;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein (HiPIP).  
 GN Hip.

OS Thiocapsa roseopersicina.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae; Thiocapsa.  
 OX NCBI\_TaxID=1058;  
 RN [1]

RP SEQUENCE.

RX MEDLINE=81094036; PubMed=7451471;

RA Tedro S.M., Meyer T.E., Bartsch R.G., Kamen M.D.;  
 RT "Primary structures of high potential, four-iron-sulfur ferredoxins  
 from the purple sulfur photosynthetic bacteria, Thiocapsa  
 roseopersicina and Chromatium gracile.";  
 RL J. Biol. Chem. 256:731-735(1981).

CC -1- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4Fe-4S

CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 CC ONE GENUS (PARACoccus) OF HALOPHILIC, DENITRIFYING BACTERIA.

CC -1- SUBUNIT: HOMODIMER (PROBABLE).

CC -1- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 CC (HIPIP) FAMILY.

DR PIR; A00264; IHTFER.

DR HSPSP; P00260; ICKU.

DR InterPro; IPR000170; Hipot\_iron\_sulf.

DR Pfam; PF01355; HiPIP\_1.

DR PRINTS; PR00374; HIPIPRDXIN.

DR PROSITE; PS00596; HiPIP; 1.

KW Electron transport; Iron-sulfur; 4Fe-4S.

FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 46 46 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 63 63 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 77 77 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

SQ SEQUENCE 85 AA; 8889 MW; C47AD747D2218482 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 DB 80 WTLK 83

## RESULT 3

Y1B1\_CLOAB

ID Y1B1\_CLOAB STANDARD; PRT; 90 AA.

AC Q04353;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CAC3711.

GN CAC3711.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=93273706; PubMed=8501044;  
 RA Sauer U., Duerre P.;  
 RT "Sequence and molecular characterization of a DNA region encoding a  
 RT small heat shock protein of Clostridium acetobutylicum.";  
 RL J. Bacteriol. 175:3394-3400(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X65276; CAA46377.1; ALT INIT.

DR EMBL; AE007866; AAK81631.1; -.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 90 AA; 10302 MW; DDDDE0F5C2FDD3CC CRC64;

Query Match 73.5%; Score 25; DB 1; Length 90;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 DB 46 WTLK 49

## RESULT 4

NULM\_BRARE

ID NULM\_BRARE STANDARD; PRT; 98 AA.

AC Q9MIY2;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).

GN MTND4L OR ND4L.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AB;

RX MEDLINE=21549095; PubMed=11691861;

RA Broughton R.E., Milam J.E., Roe B.A.;

RT "The complete sequence of the zebrafish (Danio rerio) mitochondrial

RT genome and evolutionary patterns in vertebrate mitochondrial DNA.";

RL Genome Res. 11:1958-1967(2001).

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AC024175; AAF74305.1; -  
 DR ZFIN; ZDB-GENE-011205-11; mtnd4L.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidored4L.  
 DR Pfam; PF00420; oxidored\_g2; 1.  
 DR ProDom; PD000359; Oxidored4L; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 98 AA; 10523 MW; 54C4B6230D23EF24 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 47 WTLK 50

## RESULT 5

ACHA NAJNA STANDARD; PRT; 104 AA.  
 AC P14143;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acetylcholine receptor protein, alpha chain (Fragment).  
 OS Naja naja (Indian cobra).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;  
 OC Elapidae; Elapinae; Naja.  
 OX NCBI\_TaxID=35670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89386720; PubMed=2780569;  
 RA Neumann D., Barchan D., Horowitz M., Kochva E., Fuchs S.;  
 RT "Snake acetylcholine receptors: cloning of the domain containing the  
 RT four extracellular cysteines of the alpha subunit";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7255-7259(1989).  
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA CHAINS.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M26388; AAA49384.1; -  
 DR PIR; A41384; A41384.  
 DR InterPro; IPR001175; Neur channel.  
 DR Pfam; PF02931; Neur chan LBD; 1.  
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein.  
 FT NON TER 1 1  
 FT DOMAIN <1 >104 EXTRACELLULAR.  
 FT DISULFID 10 24 BY SIMILARITY.  
 FT DISULFID 74 75 ASSOCIATED WITH RECEPTOR ACTIVATION.  
 FT CARBOHYD 23 23 N-LINKED (GLCNAC... ) (PROBABLE).  
 FT NON TER 104 104  
 SQ SEQUENCE 104 AA; 12194 MW; 47A39E2C9BFEA7A0 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 6

ACPS MYCGE STANDARD; PRT; 114 AA.  
 AC Q9ZB79;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Holo-[acyl-carrier protein] synthase (EC 2.7.8.7) (Holo-ACP synthase)  
 DE (4'-phosphopantetheinyl transferase acps).  
 GN ACPS OR MG211.1.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Transfers the 4'-phosphopantetheine moiety from coenzyme  
 CC A to a Ser of acyl-carrier protein (By similarity).  
 CC -!- CATALYTIC ACTIVITY: CoA + apo-[acyl-carrier protein] = adenosine  
 CC 3', 5'-bisphosphate + holo-[acyl-carrier protein].  
 CC -!- COFACTOR: Magnesium (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY. ACPS  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U39701; AAC71438.1; -  
 DR TIGR; MG211.1; -  
 DR InterPro; IPR002582; ACPS.  
 DR InterPro; IPR004568; Pantethn\_trn.  
 DR Pfam; PF01646; ACPS; 1.  
 DR ProDom; PD004282; ACPS; 1.  
 DR TIGRFAMs; TIGR00556; pantethn\_trn; 1.  
 KW Transferase; Lipid synthesis; Fatty acid biosynthesis; Magnesium;  
 FT METAL 8 8 MAGNESIUM (BY SIMILARITY).  
 FT METAL 58 58 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 114 AA; 13153 MW; 8992BBB560373E92 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 DB 54 WTLK 57

RESULT 7  
 HPIS\_CHRVI STANDARD; PRT; 122 AA.  
 AC P00260; P96753; Q9R4K3;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein precursor (HiPIP).  
 GN HIP.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochromatium.  
 OX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=97320627; PubMed=9177478;  
 RA Bruser T., Truper H.G., Dahl C.;  
 RT "Cloning and sequencing of the gene encoding the high potential iron-  
 sulfur protein (HiPIP) from the purple sulfur bacterium Chromatium  
 vinosum.";  
 RL Biochim. Biophys. Acta 1352:18-22(1997).  
 RN [2]  
 RP SEQUENCE OF 38-122.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=74012043; PubMed=4745771;  
 RA Dus K., Pedro S., Bartsch R.G.;  
 RT "The complete amino acid sequence of Chromatium high potential iron  
 sulfur protein.";  
 RL J. Biol. Chem. 248:7318-7331(1973).  
 RN [3]  
 RP REVISIONS TO 48; 82 AND 111.  
 RX MEDLINE=81094036; PubMed=7451471;  
 RA Tedro S.M., Meyer T.E., Bartsch R.G., Kamen M.D.;  
 RT "Primary structures of high potential, four-iron-sulfur ferredoxins  
 from the purple sulfur photosynthetic bacteria, Thiocapsa  
 roseopersicina and Chromatium gracile.";  
 RL J. Biol. Chem. 256:731-735(1981).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=74309824; PubMed=4855287;  
 RA Carter C.W. Jr., Kraut J., Freer S.T., Xuong N.H., Alden R.A.,  
 RA Bartsch R.G.;  
 RT "2-A crystal structure of oxidized Chromatium high potential iron  
 protein.";  
 RL J. Biol. Chem. 249:4212-4225(1974).  
 RN [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92135210; PubMed=1734968;  
 RA Nettesheim D.G., Harder S.R., Feinberg B.A., Otvos J.D.;  
 RT "Sequential resonance assignments of oxidized high-potential  
 iron-sulfur protein from Chromatium vinosum.";  
 RL Biochemistry 31:1234-1244(1992).  
 RN [6]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92304939; PubMed=1610810;  
 RA Gaillard J., Albrand J.-P., Moullis J.-M., Wemmer D.E.;  
 RT "Sequence-specific assignments of the 1H nuclear magnetic resonance  
 spectra of reduced high-potential ferredoxin (HiPIP) from Chromatium  
 vinosum.";  
 RL Biochemistry 31:5632-5639(1992).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95118989; PubMed=7819198;

RA Banci L., Bertini I., Dikiiy A., Kastrau D.H.W., Luchinat C.,  
 RA Sompornpisut P.;  
 RT "The three-dimensional solution structure of the reduced high-  
 potential iron-sulfur protein from Chromatium vinosum through NMR.";  
 RL Biochemistry 34:206-219(1995).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95359150; PubMed=7632685;  
 RA Bertini I., Dikiiy A., Kastrau D.H., Luchinat C., Sompornpisut P.;  
 RT "Three-dimensional solution structure of the oxidized high potential  
 iron-sulfur protein from Chromatium vinosum through NMR. Comparative  
 analysis with the solution structure of the reduced species.";  
 RL Biochemistry 34:9851-9858(1995).  
 RN [9]  
 RP STRUCTURE BY NMR OF MUTANT SER-114.  
 RX MEDLINE=96216867; PubMed=8639555;  
 RA Bontrop D., Bertini I., Cappozzi F., Dikiiy A., Eltis L., Luchinat C.;  
 RT "Three-dimensional structure of the reduced C77S mutant of the  
 Chromatium vinosum high-potential iron-sulfur protein through nuclear  
 magnetic resonance; comparison with the solution structure of the  
 wild-type protein.";  
 RL Biochemistry 35:5928-5936(1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (0.93 ANGSTROMS).  
 RX MEDLINE=20003126; PubMed=10531472;  
 RA Parisini E., Capozzi F., Lubini P., Lamzin V., Luchinat C.,  
 RA Sheldrick G.M.;  
 RT "Ab initio solution and refinement of two high-potential iron protein  
 structures at atomic resolution.";  
 RL Acta Crystallogr. D 55:1773-1784(1999).  
 CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
 FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 ONE GENUS (PARACOCUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
 CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +360 MV.  
 CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 (HIPIP) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U81381; AAB48829.1; -.  
 DR PIR; A00263; IHKREV.  
 DR PDB; HIP; 15-APR-93.  
 DR PDB; IHRQ; 03-JUN-95.  
 DR PDB; IHRR; 31-JUL-95.  
 DR PDB; INEH; 08-MAR-96.  
 DR PDB; INOE; 10-JUN-96.  
 DR PDB; IBOY; 16-DEC-98.  
 DR PDB; ICKU; 13-MAY-99.  
 DR InterPro; IPR000170; Hipot\_iron-sulf.  
 DR Pfam; PF01355; HIPIP; 1.  
 DR PRINTS; PR00374; HIPIPFDOXIN.  
 DR PROSITE; PS00596; HIPIP; 1.  
 KW Electron transport; Iron-sulfur; 4Fe-4S; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 37  
 FT CHAIN 38 122 HIGH POTENTIAL IRON-SULFUR PROTEIN.  
 FT METAL 80 80 IRON-SULFUR (4FE-4S).  
 FT METAL 83 83 IRON-SULFUR (4FE-4S).  
 FT METAL 100 100 IRON-SULFUR (4FE-4S).  
 FT METAL 114 114 IRON-SULFUR (4FE-4S).  
 FT CONFLICT 111 111 N -> D (IN REF. 3).  
 FT TURN 41 42  
 FT STRAND 43 43  
 FT TURN 46 47

RESULT 10			
ID	RS3A_ARCFU	STANDARD;	PRT; 211 AA.
AC	DC 02764;		
DT	DT 30-MAY-2000	(Rel. 39, Created)	
DT	DT 30-MAY-2000	(Rel. 35, Last sequence update)	
DT	DT 16-OCT-2001	(Rel. 40, Last annotation update)	
DE	DE 30S ribosomal protein S3Ae.		
GN	GN RPS3AE OR AF2320.		
OS	OS Archaeoglobus fulgidus.		

```

CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
CC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Klenk H.-P., Clodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Graham D.E., Kyrpides N.C.,
RA Reichmann D.L., Kerlavage A.R., Saiton G.G., Gill S.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Saiton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE S3AE FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000944; AAB88936.1; -.
DR TIGR; AF2320; -.
DR InterPro; IPR001593; Ribosomal_S3AE.
DR Pfam; PF01015; Ribosomal_S3AE; 1.
DR ProDom; PD003035; Ribosomal_S3AE; 1.
DR PROSITE; PS00191; RIBOSOMAL_S3AE; FALSE NEG.
Kw Ribosomal protein; Complete proteome.
SQ SEQUENCE 211 AA; 24334 MW; 3463C7042B0F2796 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 14 WTLK 17

RESULT 11
VE4_CRPVK
ID VE4_CRPVK STANDARD; PRT; 212 AA.
AC P03124;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Probable E4 protein.
GN E4.
OS Cottontail rabbit (shope) papillomavirus (strain Kansas) (CRPV).
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=31553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85166175; PubMed=2984661;
RA Giri I., Danos O., Yaniv M.;
RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K02708; -. NOT_ANNOTATED_CDS.
DR PIR; A03674; W4WLRB.
KW Early protein.
SQ SEQUENCE 212 AA; 24104 MW; 9FB8DA3FB14ADB5B CRC64;

Query Match 73.5%; Score 25; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 73 WTLK 76

RESULT 12
Y363_RICPR
ID Y363_RICPR STANDARD; PRT; 218 AA.
AC Q9ZDG6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP363.
GN RP363.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E.
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
CC -!- SIMILARITY: SOME TO R-PROWAZEKII RP364.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ235271; CAA14822.1; -.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 218 AA; 25780 MW; 2D8BDD7DB938F954 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 46 WTLK 49

RESULT 13
MSRA_BOVIN
ID MSRA_BOVIN STANDARD; PRT; 233 AA.
AC P54149;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide methionine sulfoxide reductase (EC 1.8.4.6) (Protein-
DE methionine-S-oxide reductase) (Peptide Met(O) reductase).

```



GN MSRA.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=99113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Adrenal medulla;  
RX MEDLINE=20519025; PubMed=8700890;  
RA Moskovitz J., Weissbach H., Brot N.;  
RT "Cloning the expression of a mammalian gene involved in the reduction  
of methionine sulfoxide residues in proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:2095-2099(1996).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
RX MEDLINE=20519025; PubMed=11063566;  
RA Lowther W.T., Brot N., Weissbach H., Matthews B.W.;  
RT "Structure and mechanism of peptide methionine sulfoxide reductase, an  
'anti-oxidation' enzyme.";  
RL Biochemistry 39:13307-13312(2000).  
CC -!- FUNCTION: HAS AN IMPORTANT FUNCTION AS A REPAIR ENZYME FOR  
PROTEINS THAT HAVE BEEN INACTIVATED BY OXIDATION. CATALYZES THE  
REVERSIBLE OXIDATION-REDUCTION OF METHIONINE SULFOXIDE IN PROTEINS  
TO METHIONINE.  
CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =  
protein L-methionine S-oxide + reduced thioredoxin.  
CC -!- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U37150; AAC48539.1; -;  
DR PDB; 1FVA; 08-NOV-00.  
DR PDB; 1FVG; 08-NOV-00.  
DR InterPro; IPR002569; PMSR.  
DR Pfam; PF01625; PMSR; 1.  
DR ProDom; PD003489; PMSR; 1.  
DR TIGRFAMs; TIGR00401; msrA; 1.  
DR Oxidoreductase; 3D-structure.  
SQ SEQUENCE 233 AA; 25846 MW; 1F69D612723FBABA CRC64;

Query Match 73.5%; Score 25; DB 1; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
Db 81 WTLK 84

RESULT 14  
PSE2\_RAT  
ID\_PSE2\_RAT STANDARD; PRT; 238 AA.  
AC Q63798;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta  
subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease  
subunit 2) (11S regulator complex beta subunit) (REG-beta).  
DE subunit 2) (11S regulator complex beta subunit) (REG-beta).  
GN PSME2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=101116;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=95309399; PubMed=7789512;  
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,  
RA Chung C.H., Shimbara N., Willy P.J., Mott J.D., Slaughter C.A.,  
RA DeMartino G.N.;  
RT "Primary structures of two homologous subunits of PA28, a gamma-  
interferon-inducible protein activator of the 20S proteasome.";  
RL FEBS Lett. 366:37-42(1995).  
CC -!- FUNCTION: Implicated in immunoproteasome assembly and required for  
efficient antigen processing. The PA28 activator complex enhances  
the cleavage pattern of the proteasome  
CC -!- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC  
RING.  
CC -!- INDUCTION: BY INTERFERON GAMMA.  
CC -!- SIMILARITY: BELONGS TO THE PA28 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D45250; BAA08207.1; -;  
DR HSSP; O06323; IAVO.  
DR InterPro; IPR003185; PA28\_alpha.  
DR InterPro; IPR003186; PA28\_beta.  
DR Pfam; PF02251; PA28\_alpha; 1.  
DR Pfam; PF02252; PA28\_beta; 1.  
KW Proteasome; Interferon induction.  
SQ SEQUENCE 238 AA; 26857 MW; AFA0F013CECE1CD3 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 238;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
Db 111 WTLK 114

RESULT 15  
PSE2\_HUMAN  
ID\_PSE2\_HUMAN STANDARD; PRT; 239 AA.  
AC Q9UL46; Q15129;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta  
subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease  
subunit 2) (11S regulator complex beta subunit) (REG-beta).  
GN PSME2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99216442; PubMed=10199920;  
RA McCusker D., Wilson M., Trowsdale J.;  
RT "Organization of the genes encoding the human proteasome activators  
PA28 alpha and PA28 beta.";  
RL Immunogenetics 49:438-445(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95309399; PubMed=7789512;  
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,  
RA Chung C.H., Shimbara N., Willy P.J., Mott J.D., Slaughter C.A.,  
RA DeMartino G.N.;  
RT "Primary structures of two homologous subunits of PA28, a gamma-  
interferon-inducible protein activator of the 20S proteasome.";  
RL FEBS Lett. 366:37-42(1995).

```

RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Implicated in immunoproteasome assembly and required for
CC efficient antigen processing. The PA28 activator complex enhances
CC the generation of class I binding peptides by altering the
CC cleavage pattern of the proteasome.
CC -!- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC
CC RING.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE PA28 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF079558; AAF02218.1; -.
DR EMBL; D45248; BAA08205.1; -.
DR EMBL; BC004368; AAH04368.1; -.
DR HSSP; Q06323; IAVO.
DR PHCI-2DPAGE; Q15129; -.
DR Genew; HGNC:9569; PSME2.
DR MIN; 602161; -.
DR InterPro; IPR003185; PA28_alpha.
DR InterPro; IPR003186; PA28_beta.
DR Pfam; PF02251; PA28_alpha; 1.
DR Pfam; PF02252; PA28_beta; 1.
KW Proteasome; Interferon induction.
FT CONFLICT 229 229 N -> T (IN REF. 2).
SQ SEQUENCE 239 AA; 27361 MW; 97A29583AA78A87C CRC64;

Query Match 73.5%; Score 25; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
DB 112 WTLK 115

```

Search completed: January 29, 2003, 10:45:07  
 Job time : 7.90909 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:40:47 ; Search time 20.8 seconds  
(without alignments)  
128.779 Million cell updates/sec

Title: US-09-707-738-23  
Perfect score: 34  
Sequence: 1 XXXXXWTLKXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	73.5	46	11 Q99LN4	Q99LN4 mus musculu
2	25	73.5	54	16 Q8VJY9	Q8VJY9 mycobacteri
3	25	73.5	60	12 Q8VAH9	Q8VAH9 white spot
4	25	73.5	64	13 Q9W7K4	Q9W7K4 naja haje.
5	25	73.5	68	2 Q8VJW6	Q8VJW6 staphylococ
6	25	73.5	70	4 Q9H1X5	Q9H1X5 homo sapien
7	25	73.5	70	16 Q8YVW8	Q8YVW8 anabaena sp
8	25	73.5	72	7 Q98260	Q98260 gadus morhu
9	25	73.5	89	12 Q919I3	Q919I3 culex nigri
10	25	73.5	92	10 Q9AR33	Q9AR33 oryza sativ
11	25	73.5	94	4 Q9HCV7	Q9HCV7 homo sapien
12	25	73.5	96	12 Q67664	Q67664 garlic late
13	25	73.5	96	15 Q9IQC1	Q9IQC1 human immun
14	25	73.5	96	15 Q8QDN8	Q8QDN8 simian-huma
15	25	73.5	99	12 Q85061	Q85061 cottontail
16	25	73.5	106	4 Q96C09	Q96C09 homo sapien

17	25	73.5	121	8 Q9G5R3	Q9G5R3 celleporell
18	25	73.5	124	16 Q988A7	Q988A7 rhizobium 1
19	25	73.5	127	16 Q97868	Q97868 clostridium
20	25	73.5	129	2 Q93SA8	Q93SA8 staphylococ
21	25	73.5	134	9 Q64107	Q64107 bacteriopho
22	25	73.5	134	16 Q34401	Q34401 bacillus su
23	25	73.5	136	16 Q8XIG7	Q8XIG7 clostridium
24	25	73.5	137	16 Q9ACR1	Q9ACR1 streptomyce
25	25	73.5	137	17 Q97UR8	Q97UR8 sulfolobus
26	25	73.5	139	6 Q9SLW6	Q9SLW6 macaca faec
27	25	73.5	139	11 Q9D8Y2	Q9D8Y2 mus musculu
28	25	73.5	143	10 Q9M673	Q9M673 cucumis mel
29	25	73.5	145	5 Q9NME6	Q9NME6 leishmania
30	25	73.5	145	16 Q8XP60	Q8XP60 clostridium
31	25	73.5	146	4 Q9H1X4	Q9H1X4 homo sapien
32	25	73.5	149	10 Q94DP9	Q94DP9 oryza sativ
33	25	73.5	150	5 Q9VQ18	Q9VQ18 dreophila
34	25	73.5	152	17 Q97YE2	Q97YE2 sulfolobus
35	25	73.5	152	17 Q97TX2	Q97TX2 sulfolobus
36	25	73.5	155	2 Q93RE2	Q93RE2 comamonas t
37	25	73.5	155	12 Q11842	Q11842 abutilon mo
38	25	73.5	158	12 Q56249	Q56249 tobacco rat
39	25	73.5	159	17 Q9VF51	Q9VF51 aeropyrum p
40	25	73.5	160	5 P91046	P91046 caenorhabdi
41	25	73.5	163	3 Q03418	Q03418 saccharomyc
42	25	73.5	163	10 Q9FHQ9	Q9FHQ9 arabidopsi
43	25	73.5	164	16 Q99X02	Q99X02 staphylococ
44	25	73.5	166	2 P71158	P71158 bacteroides
45	25	73.5	168	2 Q9KX75	Q9KX75 staphylococ

## ALIGNMENTS

RESULT 1  
Q99LN4 ID Q99LN4 PRELIMINARY; PRT; 46 AA.  
AC Q99LN4; DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 5.3 kDa protein.  
GN 4921519L13RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]\_TaxID=10090;  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002300; AAH02300.1; -.  
DR MGD; MGI:1918149; 4921519L13RIK.  
KW Hypothetical protein.  
SQ SEQUENCE 46 AA; 5337 MW; 704F20C15D045368 CRC64;  
Query Match 73.5%; Score 25; DB 11; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 7 WTLK 10  
Db 37 WTLK 40  
RESULT 2  
Q8VJY9 ID Q8VJY9 PRELIMINARY; PRT; 54 AA.  
AC Q8VJY9; DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein WTL746.1.

```

GN MT1746.1.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007036; AAK46016.1; -.
DR TIGR; MT1746; -.
KW Hypothetical protein.
SQ SEQUENCE 54 AA; 5927 MW; 22E554F875A81BAF CRC64;

Query Match 73.5%; Score 25; DB 16; Length 54;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 2 WTLK 5

RESULT 3
Q8VAH9 PRELIMINARY; PRT; 60 AA.
AC Q8VAH9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Wsv435 (WSSV494).
OS White spot syndrome virus (WSSV).
OC Viruses; unclassified viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21548311; PubMed=11689662;
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RA "Complete genome sequence of the shrimp white spot bacilliform
RT virus.";
RL J. Virol. 75:11811-11820(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RX MEDLINE=20517548; PubMed=11062040;
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase.";
RL Virology 277:100-110(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RX MEDLINE=21844071; PubMed=11853398;
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells.";
RL Virology 293:44-53(2002).
RN [5]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RA LO C.F., Kou G.-H.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF32093; AAL33437.1; -.
DR EMBL; AF440570; AAL89362.1; -.
SQ SEQUENCE 60 AA; 7091 MW; 4B4D84190256B053 CRC64;

Query Match 73.5%; Score 25; DB 12; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 31 WTLK 34

RESULT 4
Q9W7K4 PRELIMINARY; PRT; 64 AA.
AC Q9W7K4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Muscle nicotinic acetylcholine receptor alpha-subunit (Fragment).
OS ACHR-ALPHA.
OS Naja haje.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroides;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8639;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKELETAL MUSCLE;
RX MEDLINE=21396406; PubMed=11504859;
RA Takacs Z., Wilhelmsen K.C., Sorosa S.;
RT "Snake alpha-Neurotoxin Binding Site on the Egyptian Cobra (Naja haje)
RT Nicotinic Acetylcholine Receptor Is Conserved.";
RL Mol. Biol. Evol. 18:1800-1809(2001).
DR EMBL; AF077763; AAD43963.1; -.
DR InterPro; IPR000566; Lipoclin_cytFABP.
DR InterPro; IPR001175; Neur channel.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PROSITE; PS00213; LIPOCALIN, UNKNOWN_1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 64 64
SQ SEQUENCE 64 AA; 7639 MW; 92414F6F166DB2F4 CRC64;

Query Match 73.5%; Score 25; DB 13; Length 64;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 29 WTLK 32

RESULT 5
Q8VUW6 PRELIMINARY; PRT; 68 AA.
AC Q8VUW6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ORF13.
OS Staphylococcus hominis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1290;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27844;

```

RA Katayama Y., Ito T., Hiramatsu K.;  
 RT "Localization of staphylococcal cassette chromosome in methicillin  
 RL susceptible *Staphylococci*.", EMBL/GenBank/DBJ databases.  
 DR Submitted (JUN-2001) to the  
 DR EMBL; AB063171; BAB83492.1; -.  
 SQ SEQUENCE 68 AA; 7842 MW; 57892EE656D470DC CRC64;

Query Match 73.5%; Score 25; DB 2; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 65 WTLK 68

RESULT 6  
 Q9H1X5  
 ID Q9H1X5 PRELIMINARY; PRT; 70 AA.  
 AC Q9H1X5;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE BA16L21.1.3 (NADP-dependent leukotriene B4 12-hydroxydehydrogenase  
 DE (isoform 3)).  
 GN BA16L21.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laird G.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL135787; CAC22152.1; -.  
 SQ SEQUENCE 70 AA; 7983 MW; E77FF1214CF13142 CRC64;

Query Match 73.5%; Score 25; DB 4; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 7 WTLK 10

RESULT 7  
 Q8YWZ8  
 ID Q8YWZ8 PRELIMINARY; PRT; 70 AA.  
 AC Q8YWZ8;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Asl1421.  
 GN ASL1421.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213 (2001).  
 DR EMBL; AP003585; BAB73378.1; -.  
 DR InterPro; IPR001387; HTH 3.  
 DR Pfam; PF01381; HTH 3; 1.  
 DR SMART; SM00530; HTH\_XRE; 1.

KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 70 AA; 8009 MW; 764CE5F9AC236DDA CRC64;

Query Match 73.5%; Score 25; DB 16; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 17 WTLK 20

RESULT 8  
 Q98260  
 ID Q98260 PRELIMINARY; PRT; 72 AA.  
 AC Q98260;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE MHC class I antigen (Fragment).  
 GN GAMR-UA-PA1.4.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE-SPLEEN;  
 RX MEDLINE=20009526; PubMed=10541806;  
 RA Persson A.C., Stet R.J.M., Pilstroem L.;  
 RT "Characterization of MHC class I and beta2-microglobulin sequences in  
 RT Atlantic cod reveals an unusually high number of expressed class I  
 RT genes.";  
 RL Immunogenetics 50:49-59 (1999).  
 DR EMBL; AJ133450; CAB38040.1; -.  
 DR InterPro; IPR001039; MHC.I.  
 DR ProDom; PD000050; MHC.I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 72  
 SQ SEQUENCE 72 AA; 8264 MW; FB260EAB403169BE CRC64;

Query Match 73.5%; Score 25; DB 7; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 Db 69 WTLK 72

RESULT 9  
 Q91913  
 ID Q91913 PRELIMINARY; PRT; 89 AA.  
 AC Q91913;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CUN094 hypothetical protein.  
 GN CUN094.  
 OS Culex nigripalpus baculovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
 OX NCBI\_TaxID=130556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FLORIDA1997;  
 RX MEDLINE=21488685; PubMed=11602755;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Bechnel J.J., Rock D.L., Kutish G.F.;  
 RT "Genome Sequence of a Baculovirus Pathogenic for *Culex nigripalpus*.";  
 RL J. Virol. 75:11157-11165 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=FLORIDA1997;
RA Alfonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
RA Becnel J.J., Rock D.H., Kutish G.F.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF403738; AAK94172.1; -.
KW Hypothetical protein.
SQ SEQUENCE 89 AA; 10561 MW; 01F0A22D2C792541 CRC64;

Query Match          73.5%; Score 25; DB 12; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 58 WTLK 61

RESULT 10
Q9AR33 PRELIMINARY; PRT; 92 AA.
AC Q9AR33;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE P0537A05.26 protein (P0554D10.2 protein).
DE P0537A05.26 OR P0554D10.2.
CN Oryza sativa (Rice).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto I., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0537A05."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0554D10."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002971; BAB39297.1; -.
DR EMBL; AP002869; BAB39218.1; -.
SQ SEQUENCE 92 AA; 10627 MW; 30D6BD50E14804DD CRC64;

Query Match          73.5%; Score 25; DB 10; Length 92;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 24 WTLK 27

RESULT 11
Q9HCV7 PRELIMINARY; PRT; 94 AA.
AC Q9HCV7;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE DJ1153D9.3 (novel protein).
CN DJ1153D9.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Clark G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL109006; CAC03742.1; -.
SQ SEQUENCE 94 AA; 11060 MW; E0AB7D817889A06C CRC64;

Query Match          73.5%; Score 25; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 2 WTLK 5

RESULT 12
Q67664 PRELIMINARY; PRT; 96 AA.
AC Q67664;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ORF2 protein.
OS Garlic latent virus E29-6.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=143616;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi J.N., Choi Y.H., Choi Y.D., Lee J.S.;
RT "Nucleotide sequence of a cDNA for garlic mosaic virus."
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; X67134; CAA47614.1; -.
SQ SEQUENCE 96 AA; 10771 MW; A48BD84E2F7CF25A CRC64;

Query Match          73.5%; Score 25; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 7 WTLK 10

RESULT 13
Q91QC1 PRELIMINARY; PRT; 96 AA.
AC Q91QC1;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20338596; PubMed=10881687;
RA Yamada T., Iwamoto A.;
RT "Comparison of proviral accessory genes between long-term
RT nonprogressors and progressors of human immunodeficiency virus type 1
RT infection."
RL Arch. Virol. 145:1021-1027(2000).
DR EMBL; AB034523; BAA93970.1; -.
DR InterPro; IPR000012; RetroV_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVVPRVFX.
KW AIDS.
SQ SEQUENCE 96 AA; 11517 MW; 0C3C8C3A34FC8800 CRC64;

Query Match          73.5%; Score 25; DB 15; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 14

Q8QDN8 PRELIMINARY; PRT; 96 AA.  
 AC Q8QDN8;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vpr protein.  
 GN VPR.  
 OS Simian-Human immunodeficiency virus.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=57667;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RX MEDLINE=21851104; PubMed=11861859;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RT "Evolution of a human immunodeficiency virus type 1 variant with  
 RT enhanced replication in pig-tailed macaque cells by DNA shuffling.";  
 RL J. Virol. 76:2924-2935(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF465242; AAL78992.1; ..  
 SQ SEQUENCE 96 AA; 11451 MW; E41BB11BF86CDD6E CRC64;

Query Match 73.5%; Score 25; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. NO. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 15

Q85061 PRELIMINARY; PRT; 99 AA.  
 AC Q85061;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Papillomavirus sylvilagi Washington B (E2) and (E4) genes  
 DE (Fragment).  
 GN E4.  
 OS Cottontail rabbit papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RA Wu X., Brandsma J.;  
 RT "Papilloma Formation by Cottontail Rabbit Papillomavirus (CRPV)  
 RT Requires E1 and E2 Regulatory Genes in Addition to the E6 and E7  
 RT Transforming Genes.";  
 RL J. Virol. 0:0-0(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RX MEDLINE=85166175; PubMed=2984661;  
 RA Giri I., Danos O., Yaniv M.;  
 RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).

DR EMBL; U09496; AAB60542.2; .. V -> L (IN REF. 2).  
 FT CONFLICT 28 28  
 FT NON\_TER 99 99  
 SQ SEQUENCE 99 AA; 11346 MW; CE2A896101FB0BF0 CRC64;  
 Query Match 73.5%; Score 25; DB 12; Length 99;  
 Best Local Similarity 100.0%; Pred. NO. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 73 WTLK 76

Search completed: January 29, 2003, 10:46:44  
 Job time : 22.8 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:29:22 ; Search time 26.9455 Seconds  
(without alignments)  
64.288 Million cell updates/sec

Title: US-09-707-738-23

Perfect score: 34

Sequence: 1 XXXXXWTLKXXX 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID22/cgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/cgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/cgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/cgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/cgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/cgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/cgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/cgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/cgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/cgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/cgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/cgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/cgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/cgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/cgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/cgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/cgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/cgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/cgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/cgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/cgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/cgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/cgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	73.5	4	19	AAW50120 Pan DR binding pep
2	25	73.5	8	16	AA73328 Human TSH receptor
3	25	73.5	8	16	AA73329 Human TSH receptor
4	25	73.5	11	21	AA79540 Pan-DR binder pep
5	25	73.5	11	23	AAE19459 PADRE, pan-DR bind
6	25	73.5	11	23	AAU09839 Pan-DR binder pep
7	25	73.5	13	16	AA770249 Pan DR-binding pep
8	25	73.5	13	16	AA770250 Pan DR-binding pep
9	25	73.5	13	18	AAW22120 Padre (pan-DR bind
10	25	73.5	13	18	AAW22121 Padre (pan-DR bind

11	25	73.5	13	19	AAW50125 Pan DR binding pep
12	25	73.5	13	19	AAW50126 Pan DR binding pep
13	25	73.5	13	21	AA36289 Promiscuous T help
14	25	73.5	13	21	AA52558 Universal helper T
15	25	73.5	13	22	AA73644 Pan-DR binding pep
16	25	73.5	13	22	AA99709 Pan-DR-binding pep
17	25	73.5	13	22	AA99710 Pan-DR-binding pep
18	25	73.5	13	22	AA99711 Pan-DR-binding pep
19	25	73.5	13	22	AA88271 Pan-DR-binding epi
20	25	73.5	13	22	AA88271 Pan-DR-binding epi
21	25	73.5	13	22	AA20154 PADRE peptide. Sy
22	25	73.5	13	22	AA20154 PADRE peptide. Sy
23	25	73.5	13	22	AA46167 PADRE universal T
24	25	73.5	13	23	AA49066 PADRE T-cell epito
25	25	73.5	13	23	AAU0293 158p1D7 related HL
26	25	73.5	13	23	AAU0293 Pan DR epitope pep
27	25	73.5	13	23	AA34860 Pan-DR-binding epi
28	25	73.5	13	23	AA94547 Pan-DR-binding epi
29	25	73.5	13	23	AA94471 Pan-DR-binding epi
30	25	73.5	15	18	AAU10848 Helper CD4 peptide
31	25	73.5	15	23	AAU75230 Padre (pan-DR bind
32	25	73.5	18	20	AAW89677 N-terminus of NADP
33	25	73.5	19	16	AA91051 Human platelet 12-
34	25	73.5	19	20	AA91051 Amino acids 176-19
35	25	73.5	20	22	AA19635 SEQ ID NO 353 from
36	25	73.5	20	22	AA46179 Tetanus toxoid epi
37	25	73.5	20	22	AA46182 Tetanus toxoid epi
38	25	73.5	20	22	AA46204 Human APP A-beta 1
39	25	73.5	20	22	AA49078 Amyloid beta/PADRE
40	25	73.5	20	22	AA49081 Amyloid beta/PADRE
41	25	73.5	22	23	AAU10836 Human cytomagalovi
42	25	73.5	22	23	AAU10839 Human cytomagalovi
43	25	73.5	22	23	AAU10842 Human cytomagalovi
44	25	73.5	22	23	AAU10843 Human cytomagalovi
45	25	73.5	23	23	AAU10834 Human cytomagalovi
			24	22	ABB36700 Peptide #4206 enco

#### ALIGNMENTS

RESULT 1  
AAW50120  
ID AAW50120 standard; peptide; 4 AA.  
XX  
AC AAW50120;  
XX  
DT 30-JUN-1998 (first entry)  
XX  
DE Pan DR binding peptide (17).  
XX  
KW Pan DR binding peptide; antigen binding site; MHC molecule;  
KW DR locus.  
XX  
OS Synthetic.  
XX  
PN US5736142-A.  
XX  
PD 07-APR-1998.  
XX  
PF 14-SEP-1994; 94US-0305871.  
XX  
PR 14-SEP-1994; 94US-0305871.  
PR 14-SEP-1993; 93US-0121101.  
XX (CYTE-) CYTEL CORP.  
PA Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
XX WPI; 1998-239154/21.  
XX

Peptides that bind to MHC molecules of all DR alleles - inhibiting or inducing MHC Class II mediated activation of T cells

PS Disclosure; Columns 37-38; 29pp; English.

XX The present sequence, a pan DR binding peptide, is capable of

CC binding antigen binding sites on MHC molecules, which are encoded

CC by most of the alleles of a DR locus. The peptide can be used to

CC inhibit or induce MHC Class II mediated activation of T-cells or

CC helper T-cells, which themselves mediate a CTL response. The

CC peptide can be used in mammals, especially humans, to inhibit

CC T-cell-mediated events involved in allograft rejection, allergic

CC responses and autoimmunity and as a vaccine adjuvant for enhancing

CC an immune response against an administered immunogen. The peptide

CC can be used with other immunogens to treat, e.g. prostate cancer,

CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,

CC lymphoma, CMV and condyloma acuminatum.

XX

SQ Sequence 4 AA;

Query Match 73.5%; Score 25; DB 19; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10

Db 1 WTLK 4

RESULT 2

AAAR73328

ID AAR73328 standard; Peptide; 8 AA.

AC AAR73328;

XX

DT 12-DEC-1995 (first entry)

XX

DE Human TSH receptor (residues 255-262).

XX

KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;

KW antibody; affinity; detection.

XX

OS Synthetic.

XX

PN JP07089991-A.

XX

PD 04-APR-1995.

XX

PF 28-SEP-1993; 93JP-0240853.

XX

PR 28-SEP-1993; 93JP-0240853.

XX

PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.

XX

DR WPI; 1995-167251/22.

XX

XX Novel polypeptide(s) having affinity for the human TSH receptor

PT antibody - used in detection of the TSH antibody.

XX

PS Claim 1; Page 23; 54pp; Japanese.

XX

CC Peptides with affinity to human TSH (thyroid stimulating hormone)

CC receptor antibody are used for detection of the antibody. (See also

CC AAR73201-592).

XX

SQ Sequence 8 AA;

Query Match 73.5%; Score 25; DB 16; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10

Db 4 WTLK 7

RESULT 3

AAAR73329

ID AAR73329 standard; Peptide; 8 AA.

XX

AC AAR73329;

XX

DT 12-DEC-1995 (first entry)

XX

DE Human TSH receptor (residues 257-264).

XX

KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;

KW antibody; affinity; detection.

XX

OS Synthetic.

XX

PN JP07089991-A.

XX

PD 04-APR-1995.

XX

PF 28-SEP-1993; 93JP-0240853.

XX

PR 28-SEP-1993; 93JP-0240853.

XX

PA (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.

XX

DR WPI; 1995-167251/22.

XX

XX Novel polypeptide(s) having affinity for the human TSH receptor

PT antibody - used in detection of the TSH antibody.

XX

PS Claim 1; Page 23; 54pp; Japanese.

XX

CC Peptides with affinity to human TSH (thyroid stimulating hormone)

CC receptor antibody are used for detection of the antibody. (See also

CC AAR73201-592).

XX

SQ Sequence 8 AA;

Query Match 73.5%; Score 25; DB 16; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10

Db 4 WTLK 7

RESULT 4

AAAY79540

ID AAY79540 standard; Peptide; 11 AA.

XX

AC AAY79540;

XX

DT 15-AUG-2000 (first entry)

XX

DE Pan-DR binder peptide PADRE 965.10.

XX

KW Antigen presenting cell; T-lymphocyte; T-cell; immunomodulator;

KW autoimmune disease; allergy; cancer; infection; graft rejection;

KW immunotherapy; therapy; pan-DR binding peptide; PADRE; antigen.

XX

OS Synthetic.

XX

XX Key Location/Qualifiers

FT Modified-site 2 /note= "cyclohexylalanine"

XX

PN WC2000023053-A2.

XX

PD 27-APR-2000.

XX

PF 19-OCT-1999; 99WO-US24666.

XX

PR 20-OCT-1998; 98US-0105018.  
 XX (ALBA/) ALBANI S.  
 XX Albani S;  
 XX WPI; 2000-339492/29.  
 XX  
 XX New artificial antigen presenting cells useful for isolating and  
 PT expanding T cells, and modulating T cell responses for the treatment of  
 PT e.g. autoimmune diseases, allergies  
 XX  
 XX Example 16; Page 93; 179pp; English.  
 XX  
 XX This synthetic peptide, termed PADRE 965.10, is characterised as a  
 CC pan-DR binder peptide. PADRES of comparable affinity were used as  
 CC model antigens to identify human antigen-specific T cells in an  
 CC experiment designed to demonstrate that capture of T cells by  
 CC artificial antigen presenting cells (APC) is effective in  
 CC identifying polyclonal class II restricted human T cells. The  
 CC invention is directed to artificial APC, and methods of making APC  
 CC used to isolate and expand T cell populations and to modulate  
 CC the identification and isolation and antigen-specific T cells. The  
 CC methods provide for the construction of liposomes containing  
 CC MHC:peptide complexes, accessory molecules, co-stimulatory  
 CC molecules, adhesion molecules, and other molecules irrelevant to T  
 CC cell binding or modulation that are used in the binding of  
 CC artificial APC to solid support systems that may be used in the  
 CC retrieval and identification and antigen-specific T cells. Devices  
 CC and methods are provided for treating conditions that would benefit  
 CC from modulation of T cell response, e.g. autoimmune disorders  
 CC (especially type I diabetes mellitus, multiple sclerosis,  
 CC rheumatoid arthritis, dermatomyositis, juvenile rheumatoid  
 CC arthritis and uveitis), allergies, cancer, viral infections, and  
 CC graft rejection.  
 XX  
 XX Sequence 11 AA;  
 XX  
 XX Query Match 73.5%; Score 25; DB 21; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 6 WTLK 9  
 RESULT 5  
 ID AAE19459 standard; peptide; 11 AA.  
 AC AAE19459;  
 XX  
 XX 31-MAY-2002 (first entry)  
 DT  
 XX PADRE, pan-DR binder peptide.  
 DE  
 XX Human leukocyte antigen; HLA; DR-binding peptide; therapy;  
 KW stress protein; major histocompatibility complex; MHC; antiulcer;  
 KW type I diabetes; scleroderma; heat shock protein; hsp; vitiligo;  
 KW rheumatoid arthritis; lupus erythematosus; myasthenia gravis; tumour;  
 KW ulcerative colitis; infectious disease; haemostatic; nephrotropic;  
 KW polymyositis; chronic active hepatitis; primary biliary cirrhosis;  
 KW pernicious anaemia; autoimmune thyroiditis; Sjogren's syndrome; cancer;  
 KW Grave's disease; autoimmune disease; multiple sclerosis; hepatotropic;  
 KW inflammatory bowel disease; Goodpasture's syndrome; neuroprotective;  
 KW antimicrobial; immunosuppressive; dermatological; antiinflammatory;  
 KW antiinfertility; idiopathic Addison's disease; cytostatic.  
 XX  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH

FT Misc-difference 2 /label= Unknown  
 FT  
 XX WO200212286-A2.  
 PN  
 XX 14-FEB-2002.  
 PD  
 XX  
 XX 08-AUG-2001; 2001WO-USA1656.  
 XX  
 XX 09-AUG-2000; 2000US-224104P.  
 PR  
 XX 06-APR-2001; 2001US-0828574.  
 XX  
 XX (UYCA-) UNIV CALIFORNIA SAN DIEGO.  
 PA  
 XX Albani S, Prakken BJ;  
 PI  
 XX WPI; 2002-227137/28.  
 DR  
 XX  
 XX Novel human leukocyte antigen pan DR-binding peptide, useful for  
 PT treating immune mediated diseases and conditions, has a fragment of  
 PT stress protein that binds to major histocompatibility complex class II  
 PT molecules -  
 PT  
 XX Example 3; Page 46; 68pp; English.  
 PS  
 XX The invention relates to human leukocyte antigen (HLA) pan DR-binding  
 CC peptide comprising a fragment of a stress protein that binds to one or  
 CC more major histocompatibility complex (MHC) class II molecules. The  
 CC invention also relates to heat shock protein (hsp) peptides. The peptides  
 CC of the invention and thr immunomodulating composition comprising these  
 CC peptides are useful for modulating, treating or preventing an immune-  
 CC mediated disease in a mammalian subject e.g. human, having or at risk of  
 CC having a disease including autoimmune disease, multiple sclerosis (MS),  
 CC rheumatoid arthritis, lupus erythematosus, myasthenia gravis, type I  
 CC diabetes, scleroderma, ulcerative colitis, cancer (e.g. melanoma,  
 CC lymphoma, leukaemia, lung, liver, kidney, brain, bladder solid tumours,  
 CC retinoblastoma, sarcoma and connective tissue cancers) and infectious  
 CC diseases. The peptides of the invention are also useful for screening  
 CC peptides or analogues that modulate pathogenic immune response. These  
 CC peptides are useful for treating autoimmune diseases or disorders  
 CC including Sjogren's syndrome, polymyositis, chronic active hepatitis,  
 CC mixed connective tissue disease, primary biliary cirrhosis, pernicious  
 CC anaemia, autoimmune thyroiditis, vitiligo, idiopathic Addison's disease,  
 CC gluten-sensitive enteropathy, Grave's disease, inflammatory bowel  
 CC disease, autoimmune neutropaenia, idiopathic thrombocytopenia purpura,  
 CC pemphigus vulgaris, autoimmune infertility, Goodpasture's syndrome,  
 CC bullous pemphigoid, discoid lupus and dense deposit disease. The present  
 CC sequence is PADRE, a pan-DR binder peptide used in the exemplification  
 CC of the invention.  
 XX  
 XX Sequence 11 AA;  
 XX  
 XX Query Match 73.5%; Score 25; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 6 WTLK 9  
 RESULT 6  
 ID AAU09839 standard; peptide; 11 AA.  
 XX  
 XX AAU09839;  
 AC  
 XX 26-FEB-2002 (first entry)  
 DT  
 XX Pan-DR binder peptide (PADRE).  
 DE  
 XX Antigenic; antidiabetic; neuroprotective; antirheumatic; antiarthritic;  
 KW dermatological; immunosuppressive; ophthalmological; antiallergic;

KW cytostatic; virucide; antibacterial; artificial antigen presenting cell;  
KW liposome; major histocompatibility complex; MHC; T cell; allograft;  
KW cytokine production; T-helper 2 response; transplantation therapy;  
KW autoimmune disease; type I diabetes mellitus; multiple sclerosis;  
KW rheumatoid arthritis; dermatomyositis; juvenile rheumatoid arthritis;  
KW uveitis; allergy; cancer; infection; Pan-DR binder peptide; PADRE.  
XX  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /label= OTHER  
FT /note= cyclohexylalanine"  
XX  
XX WO2001180833-A1.  
XX  
XX 01-NOV-2001.  
XX  
XX 20-APR-2000; 2000WO-IT00161.  
XX  
XX 20-APR-2000; 2000WO-IT00161.  
XX  
XX (ALBA/) ALBANI S.  
XX  
XX Albani S;  
XX  
XX WPI; 2002-055316/07.  
XX  
XX New artificial antigen presenting cell, useful for modulating T cell  
PT response for treating allergies and cancers, comprises liposome, major  
PT histocompatibility complex, antigen and accessory molecule components -  
PT  
XX Example 16; Page 103; 195pp; English.  
XX  
XX The invention relates to an artificial antigen presenting cell (1)  
CC comprising liposome, major histocompatibility complex (MHC), antigen and  
CC accessory molecule components. (1) is used in methods for the  
CC following: (1) identifying T cells specific for an antigen of interest;  
CC (2) isolating T cells specific for an antigen of interest; (3) modulating  
CC T cell response; (4) characterising the functional state of antigen-  
CC specific T cells; (5) treating a condition in a subject which would be  
CC benefited by altering the functional pattern of cytokine production by  
CC certain antigen-specific T cells to increase T-helper (Th) 2 response  
CC and/or decrease Th1 response; (6) identifying antigen-specific T cells  
CC specific for epitopes on a graft donor's tissue likely to elicit graft  
CC versus host rejection response; and (7) treating a recipient mammal to  
CC reduce rejection of allografts in a transplantation therapy regime.  
CC Method (5) is useful for treating autoimmune disease such as type I  
CC diabetes mellitus, multiple sclerosis, rheumatoid arthritis,  
CC dermatomyositis, juvenile rheumatoid arthritis or uveitis. Alternatively  
CC it is useful for treating allergy due to allergens such as dust, animal  
CC skin bypass products, vegetables, fruits, pollen or chemicals, cancer,  
CC viral infection, bacterial infection. Addition of the accessory  
CC molecules, as well as co-stimulatory molecules, and other proteins in  
CC proper orientation in the liposomes allow for substantially improved  
CC binding association and manipulation of T cells which is very important  
CC in the identification and stimulation of antigen-specific T cells. The  
CC present sequence represents Pan-DR binder peptide (PADRE) used in  
CC the method of the invention.  
XX  
XX  
SQ Sequence 11 AA;  
Query Match 73.5%; Score 25; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 WTLK 10  
|||||  
Db 6 WTLK 9  
|||||  
RESULT 7  
AAR70249

ID AAR70249 standard; peptide; 13 AA.  
XX  
AC AAR70249;  
XX  
DT 13-NOV-1995 (first entry)  
XX  
XX Pan DR-binding peptide activator of T cells.  
XX  
XX T cell; MHC; class II; major histocompatibility complex II;  
KW immunogen; activation.  
KW  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 1 /note= "D-form residue"  
FT  
FT Modified-site 3 /label= OTHER  
FT /note= "cyclohexylalanine"  
FT Misc-difference 13 /note= "D-form residue"  
FT  
XX WO9507707-A.  
XX  
XX 23-MAR-1995.  
XX  
XX 14-SEP-1994; 94WO-US10368.  
XX  
XX 14-SEP-1993; 93US-0121101.  
XX  
XX (CYTE-) CYTEL CORP.  
XX  
XX Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
XX  
XX WPI; 1995-131178/17.  
XX  
XX Inhibiting or inducing an immune response using Pan DR-binding  
PT peptide(s) - to prepare a compsn. used to treat allo:graft  
PT rejection, allergic response and auto:immunity and as a vaccine  
PT component  
XX  
XX Claim 3; Page 51; 59pp; English.  
XX  
XX Peptides capable of binding MHC class II molecules via antigen  
CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
CC AAR70249 is capable of binding MHC class II molecules causing T cell  
CC activation. This is useful as a vaccine component and may be used in  
CC conjunction with CTL peptides to enhance a CTL response.  
XX  
XX Sequence 13 AA;  
SQ  
Query Match 73.5%; Score 25; DB 16; Length 13;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 WTLK 10  
|||||  
Db 7 WTLK 10  
|||||  
RESULT 8  
AAR70250  
ID AAR70250 standard; peptide; 13 AA.  
XX  
XX AAR70250;  
XX  
XX 13-NOV-1995 (first entry)  
XX  
XX Pan DR-binding peptide activator of T cells.  
XX  
XX T cell; MHC; class II; major histocompatibility complex II;  
KW immunogen; activation.  
KW  
XX

OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 FT  
 XX WO9507707-A.  
 XX  
 XX 23-MAR-1995.  
 XX  
 XX 14-SEP-1994; 94WO-US10368.  
 XX  
 XX 14-SEP-1993; 93US-0121101.  
 XX  
 XX (CYTE-) CYTEL CORP.  
 PA  
 PA Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;  
 PI WPI, 1995-131178/17.  
 XX  
 XX Inhibiting or inducing an immune response using Pan DR-binding  
 PT peptide(s) - to prepare a compsn. used to treat allo:graft  
 PT rejection, allergic response and auto:immunity and as a vaccine  
 PT component  
 XX  
 PS Claim 3; Page 51; 59pp; English.  
 XX  
 CC Peptides capable of binding MHC class II molecules via antigen  
 CC binding sites (pan DR-binding peptides) were synthesised. Peptide  
 CC AAR70250 is capable of binding MHC class II molecules causing T cell  
 CC activation. This can be used as a vaccine component. The peptide can also  
 CC be used in combination with CTL peptides to enhance a CTL response.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 73.5%; Score 25; DB 16; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 Db |||||  
 7 WTLK 10  
 Db  
 RESULT 9  
 AAW22120  
 ID AAW22120 standard; peptide; 13 AA.  
 XX  
 AC AAW22120;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 5.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Modified-site 3 /note= "cyclohexylalanine"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 FT  
 XX WO9726784-A1.  
 XX  
 XX 31-JUL-1997.  
 PD  
 XX

PF 23-JAN-1997; 97WO-US01041.  
 XX  
 DR 24-JAN-1996; 96US-0010510.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 XX  
 DR Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 PT prevention and treatment of tumours and infections  
 XX  
 XX Claim 23; Page 74; 87pp; English.  
 XX  
 CC This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 73.5%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 Db |||||  
 7 WTLK 10  
 Db  
 RESULT 10  
 AAW22121  
 ID AAW22121 standard; peptide; 13 AA.  
 XX  
 AC AAW22121;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Padre (pan-DR binding) peptide 6.  
 XX  
 KW PADRE; pan-DR binding; immune response; antigenic determinant;  
 KW treatment; tumour; infection.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT  
 FT  
 XX WO9726784-A1.  
 XX  
 XX 31-JUL-1997.  
 PD  
 XX  
 XX 23-JAN-1997; 97WO-US01041.  
 XX  
 XX 24-JAN-1996; 96US-0010510.  
 PR  
 XX

PA (CYTE-) CYTEL CORP.  
 XX Alexander JL, Defrees S, Sette A;  
 XX WPI; 1997-393272/36.  
 XX  
 XX Composition for eliciting immune response to non-protein determinant  
 PT - comprises the determinant and a pan-DR binding peptide, used for  
 PT prevention and treatment of tumours and infections  
 XX  
 XX Claim 23; Page 74; 87pp; English.  
 XX  
 XX This PADRE (pan-DR binding) peptide is used in a composition for  
 CC eliciting an immune response to a non-protein antigenic determinant.  
 CC The composition comprises of the PADRE peptide covalently linked to the  
 CC non-protein antigenic determinant. The composition is used to induce a  
 CC therapeutic or prophylactic response, particularly to selected  
 CC polysaccharide antigens associated with tumours or infectious agents. It  
 CC provides a high level, long-lasting IgG immune response. The composition  
 CC is also used to produce monoclonal antibodies which are potentially  
 CC useful as therapeutic and diagnostic agents. The composition can be used  
 CC to diagnose susceptibility of a patient to treatment with the non-protein  
 CC antigenic determinant or to predict subjects at risk from developing  
 CC chronic infections. PADRE peptides have broader specificity and higher  
 CC affinity than known DR-binding peptides. They are powerful inhibitors of  
 CC the proliferative response of human T cells restricted by at least 6  
 CC different DR molecules, and act as helper epitopes of in vivo induction  
 CC of cytotoxic T cells and antibody production.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 73.5%; Score 25; DB 18; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 7 WTLK 10  
 RESULT 11  
 AAW50125  
 ID AAW50125 standard; peptide; 13 AA.  
 AC AAW50125;  
 DT 30-JUN-1998 (first entry)  
 DE Pan DR binding peptide (22).  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Misc-difference 3 /label= Tyr, Phe  
 XX  
 XX US5736142-A.  
 XX 07-APR-1998.  
 XX 14-SEP-1994; 94US-0305871.  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 XX (CYTE-) CYTEL CORP.  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.

XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 39-40; 29pp; English.  
 XX  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or  
 CC helper T-cells, which themselves mediate a CTL response. The  
 CC peptide can be used in mammals, especially humans, to inhibit  
 CC T-cell-mediated events involved in allograft rejection, allergic  
 CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
 CC an immune response against an administered immunogen. The peptide  
 CC can be used with other immunogens to treat, e.g. prostate cancer,  
 CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
 CC lymphoma, CMV and condyloma acuminatum.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 73.5%; Score 25; DB 19; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 7 WTLK 10  
 RESULT 12  
 AAW50126  
 ID AAW50126 standard; peptide; 13 AA.  
 AC AAW50126;  
 DT 30-JUN-1998 (first entry)  
 DE Pan DR binding peptide (23).  
 XX Pan DR binding peptide; antigen binding site; MHC molecule;  
 KW DR locus.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Misc-difference 3 /label= Tyr, Phe  
 XX  
 XX US5736142-A.  
 XX 07-APR-1998.  
 XX 14-SEP-1994; 94US-0305871.  
 XX 14-SEP-1994; 94US-0305871.  
 PR 14-SEP-1993; 93US-0121101.  
 XX (CYTE-) CYTEL CORP.  
 XX Alexander JL, Gaeta F, Grey HM, Sette A, Sidney J;  
 XX WPI; 1998-239154/21.  
 XX Peptides that bind to MHC molecules of all DR alleles - inhibiting  
 PT or inducing MHC Class II mediated activation of T cells  
 XX  
 XX Claim 35; Columns 41-42; 29pp; English.  
 XX  
 XX The present sequence, a pan DR binding peptide, is capable of  
 CC binding antigen binding sites on MHC molecules, which are encoded  
 CC by most of the alleles of a DR locus. The peptide can be used to  
 CC inhibit or induce MHC Class II mediated activation of T-cells or

CC helper T-cells, which themselves mediate a CTL response. The  
 CC peptide can be used in mammals, especially humans, to inhibit  
 CC T-cell-mediated events involved in allograft rejection, allergic  
 CC responses and autoimmunity and as a vaccine adjuvant for enhancing  
 CC an immune response against an administered immunogen. The peptide  
 CC can be used with other immunogens to treat, e.g. prostate cancer,  
 CC hepatitis B, hepatitis C, AIDS, renal and cervical carcinoma,  
 CC lymphoma, CMV and condyloma acuminatum.

XX SQ Sequence 13 AA;

Query Match 73.5%; Score 25; DB 19; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10  
 DB 7 WTLK 10

RESULT 13

AAB36289  
 ID AAB36289 standard; Peptide; 13 AA.

XX AC AAB36289;

XX DT 26-FEB-2001 (first entry)

XX DE Promiscuous T helper epitope SEQ ID NO: 65.

XX KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;  
 KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX OS Synthetic.

XX PN WO200065058-A1.

XX PD 02-NOV-2000.

XX PF 19-APR-2000; 2000WO-DK00205.

XX PR 23-APR-1999; 99DK-0000552.

XX PR 06-MAY-1999; 99US-0132811.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Klysner S;

XX DR WPI; 2000-672791/65.

XX PT Down-regulating interleukin 5 (IL-5) activity in humans by  
 PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,  
 PT prophylaxis or amelioration of asthma or other chronic allergic  
 PT conditions -

XX PS Disclosure; Page 168; 172pp; English.

XX CC The present invention is concerned with methods of treating asthma,  
 CC eosinophilia, allergic rhinitis and other allergic diseases. These,  
 CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
 CC proteins and their coding sequences to down-regulate IL-5 activity and  
 CC thus reduce eosinophil numbers. The allergic diseases may be treated  
 CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
 CC it is possible that they may be used in the treatment of cancer and  
 CC helminthic infections.

XX SQ Sequence 13 AA;

Query Match 73.5%; Score 25; DB 21; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WTLK 10

Db 7 WTLK 10  
 RESULT 14  
 AAY52558  
 ID AAY52558 standard; peptide; 13 AA.  
 XX AC AAY52558;  
 XX DT 28-FEB-2000 (first entry)  
 XX DE Universal helper T epitope, pan DR epitope (PADRE).  
 XX KW Chimeric; Ii protein; pan DR epitope; expression vector;  
 KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
 KW epitope; antigen; presentation; class I; cytosolic pathway;  
 KW endoplasmic reticulum; class II; extracellular antigen;  
 KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;  
 KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
 KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
 KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
 KW autoimmune disease; activation; antiviral; antimalarial;  
 KW immunoprotective.  
 XX OS Synthetic.  
 XX PN WO9958658-A2.  
 XX PD 18-NOV-1999.  
 XX PF 13-MAY-1999; 99WO-US10646.  
 XX PR 13-MAY-1998; 98US-0078904.  
 XX PR 15-MAY-1998; 98US-0085751.  
 XX PA (EPIM-) EPIMMUNE INC.  
 XX PI Fikes JD, Hermanson GG, Sette A, Ishioka GV, Livingston B;  
 PI Chesnut RW;  
 XX WPI: 2000-039103/03.  
 XX DR N-PSDB; AAZ38679.  
 XX PT Expression vectors encoding major histocompatibility targeting  
 PT sequence, used as, e.g. tumor vaccines -  
 XX PS Claim 9; Page 80; 130pp; English.  
 XX CC This sequence represents a universal helper T epitope, pan DR epitope  
 CC (PADRE), DNA encoding which is used to construct fusion genes used  
 CC in exemplifications of the present invention. The invention  
 CC relates to a novel expression vector comprising a promoter operably  
 CC linked to a fusion gene encoding a major histocompatibility complex  
 CC (MHC) targeting sequence, and two or more heterologous peptide epitopes.  
 CC The MHC targeting sequence may be a class I targeting sequence, which  
 CC directs an MHC class I epitope to a cytosolic pathway or to the  
 CC endoplasmic reticulum, or an MHC class II targeting sequence, which  
 CC directs extracellular antigens to enter the endocytic pathway to be  
 CC processed into antigen peptides for presentation on MHC class II  
 CC molecules. The heterologous epitopes may comprise either helper T  
 CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and  
 CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are  
 CC useful for stimulating an immune response in vivo, as well as for use in  
 CC assaying the human immunogenicity of a human T cell peptide epitope in  
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
 CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
 CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
 CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
 CC combined with other MHC class I and class II epitopes to increase the  
 CC number of cells that are activated in response to a given antigen and  
 CC provide a broader population coverage of MHC-reactive alleles.

XX SQ Sequence 13 AA;  
 Query Match 73.5%; Score 25; DB 21; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 Db 7 WTLK 10

RESULT 15  
 AAB73644  
 ID AAB73644 standard; peptide; 13 AA.  
 XX  
 AC AAB73644;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Pan-DR binding peptide, PADRE.  
 XX  
 KW Cellular vaccine; antigen-presenting cell; APC;  
 KW cell-surface molecule density; major histocompatibility complex; MHC;  
 KW antigen-specific T-cell expansion; tumour; cancer; viral infection;  
 KW parasitic infection; Pan-DR binding peptide; PADRE.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200136978-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 15-NOV-2000; 2000WO-US42213.  
 XX  
 PR 15-NOV-1999; 99US-0165428.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schultze JL, Vonderheide RH, Nadler LM, Maecker B;  
 PI Von Bergwelt-Baildon M;  
 XX  
 XX WPI; 2001-343909/36.  
 XX  
 XX Determining surface density of molecules on antigen-presenting cells,  
 PT useful for quality control of cellular vaccines for treatment of tumors  
 PT and infections comprises measuring density ratios -  
 XX  
 PS Disclosure; Page 21; 47pp; English.  
 XX  
 CC The invention relates to a method of determining the surface density of  
 CC a cell-surface molecule on a primary or artificial antigen-presenting  
 CC cell (APC). The method comprises determining the cell surface area  
 CC of the APCs, determining the absolute amount of cell surface molecule on  
 CC the surface of the APCs, and calculating the ratio of the amount of  
 CC the cell surface molecule to the APC surface area as a measure of cell  
 CC surface molecule density. The invention also encompasses a kit for  
 CC determining the cell-surface density of a complex of peptide and MHC  
 CC (major histocompatibility complex) protein/peptide complex on an APC;  
 CC identifying an agent that increases persistence of the MHC/peptide  
 CC complex on an APC surface by culturing APCs in presence of test compound  
 CC and measuring the time of persistence relative to an untreated control;  
 CC and determining if the APC carries a therapeutically adequate amount of  
 CC peptide by determining whether the cell-surface density of the  
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The  
 CC method is used to determine if the amount of MHC/antigenic peptide  
 CC complex present on the surface of an APC is sufficient for therapeutic  
 CC use of the cells as cellular vaccines or for ex vivo expansion of  
 CC antigen-specific T cells for subsequent return to the patient,  
 CC particularly for the treatment of tumours or viral or parasitic  
 CC infections. The method can also be used to identify agents that increase  
 CC production of MHC/peptide complexes on cells, (which increases the  
 CC therapeutic potential of the cells) by treatment before and/or during

CC APC administration. The method allows identification of therapeutically  
 CC useful antigen-pulsed APC optimisation of conditions for their  
 CC production, and control of APC quality. Sequences AAB73643-AAB73648  
 CC represent peptides used in MHC binding studies and for the generation  
 CC of peptide-specific cytotoxic T-lymphocytes.

XX SQ Sequence 13 AA;

Query Match 73.5%; Score 25; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 Db 7 WTLK 10

Search completed: January 29, 2003, 10:44:34  
 Job time : 26.9455 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 29, 2003, 10:46:52 ; Search time 5.67273 Seconds  
(without alignments)  
46.243 Million cell updates/sec

Title: US-09-707-738-23

Perfect score: 34

Sequence: 1 XXXXXWTLKXXX 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	73.5	11	10	US-09-756-983-7
2	25	73.5	11	10	US-09-828-574-12
3	25	73.5	13	9	US-09-949-375A-18
4	25	73.5	13	9	US-09-785-215-19
5	25	73.5	13	10	US-09-894-018-69
6	25	73.5	24	10	US-09-864-761-47626
7	25	73.5	24	10	US-09-983-019-10
8	25	73.5	27	10	US-09-864-761-40242
9	25	73.5	27	10	US-09-983-019-4
10	25	73.5	27	10	US-09-983-019-7
11	25	73.5	28	10	US-09-864-761-39731
12	25	73.5	48	10	US-09-864-761-45303
13	25	73.5	51	9	US-10-016-634A-165
14	25	73.5	55	9	US-09-949-375A-12
15	25	73.5	56	10	US-09-864-761-41804
16	25	73.5	57	10	US-09-864-761-44535
17	25	73.5	58	10	US-09-925-300-1859
18	25	73.5	59	10	US-09-925-299-822
19	25	73.5	62	10	US-09-896-578-5

20	25	73.5	73	10	US-09-864-761-47212	Sequence 47212, A
21	25	73.5	74	10	US-09-764-877-1943	Sequence 1943, Ap
22	25	73.5	80	10	US-09-894-018-97	Sequence 97, Appl
23	25	73.5	87	10	US-09-925-297-562	Sequence 562, App
24	25	73.5	89	10	US-09-867-550-686	Sequence 686, App
25	25	73.5	97	10	US-09-864-761-43929	Sequence 43929, A
26	25	73.5	98	10	US-09-894-018-101	Sequence 101, App
27	25	73.5	106	9	US-10-138-516-3	Sequence 3, Appli
28	25	73.5	106	9	US-10-146-130-5	Sequence 5, Appli
29	25	73.5	106	10	US-09-894-018-95	Sequence 95, Appl
30	25	73.5	107	10	US-09-894-018-93	Sequence 93, Appl
31	25	73.5	107	10	US-09-894-018-103	Sequence 103, App
32	25	73.5	109	10	US-09-910-150-32	Sequence 32, Appl
33	25	73.5	113	9	US-09-789-054A-18	Sequence 18, Appl
34	25	73.5	123	10	US-09-894-018-109	Sequence 109, App
35	25	73.5	130	10	US-09-894-018-99	Sequence 99, Appl
36	25	73.5	144	10	US-09-894-018-129	Sequence 129, App
37	25	73.5	147	10	US-09-894-018-131	Sequence 131, App
38	25	73.5	148	10	US-09-894-018-127	Sequence 127, App
39	25	73.5	152	10	US-09-939-980-264	Sequence 264, App
40	25	73.5	157	10	US-09-894-018-117	Sequence 117, App
41	25	73.5	168	10	US-09-894-018-115	Sequence 115, App
42	25	73.5	169	10	US-09-894-018-119	Sequence 119, App
43	25	73.5	170	9	US-09-479-040-3	Sequence 3, Appli
44	25	73.5	180	10	US-09-894-018-125	Sequence 125, App
45	25	73.5	183	9	US-09-893-152-19	Sequence 19, Appl

## ALIGNMENTS

RESULT 1  
US-09-756-983-7  
; Sequence 7, Application US/09756983  
; Patent No. US20020122818A1  
; GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthesized peptide totally artificial  
; OTHER INFORMATION: Xaa in position 2 stands for cyclohexylalanine  
US-09-756-983-7

Query Match 73.5%; Score 25; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||  
Db 6 WTLK 9

RESULT 2  
US-09-828-574-12  
; Sequence 12, Application US/09828574  
; Patent No. US20020146759A1

```
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: ALBANI, Salvatore
; APPLICANT: PRAKKEN, Berent J.
; TITLE OF INVENTION: STRESS PROTEINS AND PEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: UCSD1310-1
; CURRENT APPLICATION NUMBER: US/09/828,574
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/224,104
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: pan-DR binder peptide
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is any amino acid
US-09-828-574-12

Query Match          73.5%; Score 25; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 6 WTLK 9

RESULT 3
US-09-949-375A-18
; Sequence 18, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.
US-09-949-375A-18

Query Match          73.5%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 4
US-09-785-215-19
; Sequence 19, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 19
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial T-cell epitope capable of binding to a large portion
; OTHER INFORMATION: of MHC Class II molecules in a variety of animals
US-09-785-215-19

Query Match          73.5%; Score 25; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 5
US-09-894-018-69
; Sequence 69, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69

Query Match          73.5%; Score 25; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
   ||||
Db 7 WTLK 10

RESULT 6
US-09-864-761-47626
; Sequence 47626, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
```

```
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47626
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011502.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: AI033706.1, EVALUATE 1.00e-08
; OTHER INFORMATION: SWISSPROT HIT: Q84291, EVALUATE 4.90e+00
US-09-864-761-47626

Query Match 73.5%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 13 WTLK 16

RESULT 7
US-09-983-019-10
; Sequence 10, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (3)..(3)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-10

Query Match 73.5%; Score 25; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 7 WTLK 10

RESULT 8
US-09-864-761-40242
; Sequence 40242, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40242
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC025746.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-09-864-761-40242
```

```
Query Match 73.5%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 WTLK 10
|||
Db 1 WTLK 4
```

## RESULT 9

```
US-09-983-019-4
; Sequence 4, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP19
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (6)..(6)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-4
```

```
Query Match 73.5%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 WTLK 10
.||||
Db 10 WTLK 13
```

## RESULT 10

```
US-09-983-019-7
; Sequence 7, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP19
; FILE REFERENCE: 1954-347US
```

```
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: VARIANT
; LOCATION: (17)..(17)
; OTHER INFORMATION: X = cyclohexylalanine or phenylalanine
US-09-983-019-7
```

```
Query Match 73.5%; Score 25; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 WTLK 10
|||
Db 21 WTLK 24
```

## RESULT 11

```
US-09-864-761-39731
; Sequence 39731, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 39731  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC003681.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8  
US-09-864-761-39731

Query Match 73.5%; Score 25; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
DB 11 WTLK 14

## RESULT 12

US-09-864-761-45303  
; Sequence 45303, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45303  
; LENGTH: 48  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC022216.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.71  
; OTHER INFORMATION: SWISSPROT HIT: Q9ZER9, EVALUATE 1.20e+00  
US-09-864-761-45303

Query Match 73.5%; Score 25; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
DB 2 WTLK 5

## RESULT 13

US-10-016-634A-165  
; Sequence 165, Application US/10016634A  
; Publication No. US20020192666A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Ghosh, Malavika  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Prot  
; FILE REFERENCE: DEX-0255  
; CURRENT APPLICATION NUMBER: US/10/016,634A  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US 60/244,258  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 176  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 165  
; LENGTH: 51  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-016-634A-165

Query Match 73.5%; Score 25; DB 9; Length 51;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||||  
DB 41 WTLK 44

## RESULT 14

US-09-949-375A-12  
; Sequence 12, Application US/09949375A  
; Patent No. US20020172673A1  
; GENERAL INFORMATION:  
; APPLICANT: KLYSNER, Steen et al.  
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE

```
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.
US-09-949-375A-12

Query Match      73.5%; Score 25; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 WTLK 10
        ||||
Db      19 WTLK 22

RESULT 15
US-09-864-761-41804
; Sequence 41804, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41804
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010582.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.9
; OTHER INFORMATION: EST HUMAN HIT: AW873461.1, EVALUE 3.00e-18
; OTHER INFORMATION: SWISSPROT HIT: P16473, EVALUE 1.00e-27
US-09-864-761-41804
```

```
Query Match      73.5%; Score 25; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 WTLK 10
        ||||
Db      21 WTLK 24
```

```
Search completed: January 29, 2003, 10:59:00
Job time : 5.67273 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:40:47 ; Search time 22.4 Seconds  
(without alignments)  
128.779 Million cell updates/sec

Title: US-09-707-738-24  
Perfect score: 35  
Sequence: 1 XXXXXXWTLKXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organalle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rdent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	25	71.4	46	11 Q99LN4	Q99LN4 mus musculus
2	25	71.4	54	16 Q8VJY9	Q8VJY9 mycobacteri
3	25	71.4	60	12 Q8VAH9	Q8VAH9 white spot
4	25	71.4	64	13 Q9W7K4	Q9W7K4 najja hajje
5	25	71.4	68	2 Q8VUM6	Q8VUM6 staphylococ
6	25	71.4	70	4 Q9HLX5	Q9HLX5 homo sapien
7	25	71.4	70	16 Q8YWZ8	Q8YWZ8 anabaena sp
8	25	71.4	72	7 Q8Z60	Q8Z60 gadus morhu
9	25	71.4	89	12 Q919I3	Q919I3 culex nigri
10	25	71.4	92	10 Q9AR33	Q9AR33 oryza sativ
11	25	71.4	94	4 Q9HCV7	Q9HCV7 homo sapien
12	25	71.4	96	12 Q67664	Q67664 garlic late
13	25	71.4	96	15 Q91QC1	Q91QC1 human immun
14	25	71.4	96	15 Q8QDN8	Q8QDN8 sinian-huma
15	25	71.4	99	12 Q8S061	Q8S061 cotton-tall
16	25	71.4	106	4 Q96C09	Q96C09 homo sapien

## SUMMARIES

RESULT 1  
Q99LN4  
ID Q99LN4 PRELIMINARY; PRT; 46 AA.  
AC Q99LN4;  
DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)  
DE Hypothetical 5.3 kDa protein.  
GN 4921519L13RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002300; AAH02300.1; -;  
DR MGD; MGI:1918149; 4921519L13RIK.  
KW Hypothetical protein.  
SQ SEQUENCE 46 AA; 5337 MW; 704F20C15D045368 CRC64;  
Query Match 71.4%; Score 25; DB 11; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 WTLK 10  
Db 37 WTLK 40  
RESULT 2  
Q8VJY9  
ID Q8VJY9 PRELIMINARY; PRT; 54 AA.  
AC Q8VJY9;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
DE Hypothetical protein MT1746.1.

Q95r3 celloporell  
Q98a7 thizobium l  
Q97e8 clostridium  
Q93a8 staphylococ  
Q64107 bacterioph  
Q34401 bacillus su  
Q8xig7 clostridium  
Q9acr1 streptomyc  
Q97ur8 sulfolobus  
Q95lw6 macaca fasc  
Q9d8y2 mus musculu  
Q9m673 cucumis mel  
Q9nme6 leishmania  
Q8xp60 clostridium  
Q9h1x4 homo sapien  
Q94dp9 oryza sativ  
Q9vq18 drosophila  
Q97ye2 sulfolobus  
Q97tx2 sulfolobus  
Q93re2 comamonas t  
O11842 abutilon mo  
O56249 tobacco rat  
Q9yfs1 aeropyrum p  
P91046 caenorhabdi  
Q03418 saccharomyc  
Q9fhq9 arabidopsis  
Q99x02 staphylococ  
P71158 bacteroides  
Q9kx75 staphylococ

## ALIGNMENTS

GN MT1746.1.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae, Mycobacteriaceae, Mycobacterium.  
 OK NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / OSHKOSH;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE007036; AAK46016.1; -.  
 DR TIGR; MT1746; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 54 AA; 5927 MW; 22E554F875A81BAF CRC64;  
 Query Match 71.4%; Score 25; DB 16; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 2 WTLK 5  
 RESULT 3  
 Q8VAH9  
 ID Q8VAH9 PRELIMINARY; PRT; 60 AA.  
 AC Q8VAH9  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Wav435 (WSSV494).  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 OK NCBI\_TaxID=92652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21548311; PubMed=11699662;  
 RA Yang F., He J., Lin X., Li Q., Fan D., Zhang X., Xu X.;  
 RT "Complete genome sequence of the shrimp white spot bacilliform  
 RT virus."  
 RL J. Virol. 75:11811-11820(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TAIWAN;  
 RX MEDLINE=20517548; PubMed=11062040;  
 RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
 RT "Identification and characterization of a shrimp white spot syndrome  
 RT virus (WSSV) gene that encodes a novel chimeric polypeptide of  
 RT cellular-type thymidine kinase and thymidylate kinase."  
 RL Virology 277:100-110(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TAIWAN;  
 RX MEDLINE=21844071; PubMed=11853398;  
 RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
 RA Lo C.F., Kou G.H.;  
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
 RT spot syndrome virus and characterization of the motif important for  
 RT targeting VP35 to the nuclei of transfected insect cells."  
 RL Virology 293:44-53(2002).  
 RN [5]

RP SEQUENCE FROM N.A.  
 RC STRAIN=TAIWAN;  
 RA Lo C.-F., Kou G.-H.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF332093; AAL33437.1; -.  
 DR EMBL; AF440570; AAL89362.1; -.  
 SQ SEQUENCE 60 AA; 7091 MW; 4B4D84190256B053 CRC64;  
 Query Match 71.4%; Score 25; DB 12; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 31 WTLK 34  
 RESULT 4  
 Q9W7K4  
 ID Q9W7K4 PRELIMINARY; PRT; 64 AA.  
 AC Q9W7K4  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE MuscLE nicotinic acetylcholine receptor alpha-subunit (Fragment).  
 GN ACHR-ALPHA.  
 OS Naja haje.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Elapidae; Elapinae; Naja.  
 OK NCBI\_TaxID=8639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKELETAL MUSCLE;  
 RX MEDLINE=21396406; PubMed=11504859;  
 RA Takacs Z., Wilhelmsen K.C., Sorota S.;  
 RT "Snake alpha-Neurotoxin Binding Site on the Egyptian Cobra (Naja haje)  
 RT Nicotinic Acetylcholine Receptor Is Conserved."  
 RL Mol. Biol. Evol. 18:1800-1809(2001).  
 DR EMBL; AF077763; AAD43963.1; -.  
 DR InterPro; IPR000566; Lipoclin\_cytFABP.  
 DR InterPro; IPR001175; Neur channel.  
 DR Pfam; PF02931; Neur chan IBD; 1.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
 KW Receptor.  
 FT NON\_TER 1 64  
 FT NON\_TER 64 64  
 SQ SEQUENCE 64 AA; 7639 MW; 92414F6F166DB2F4 CRC64;  
 Query Match 71.4%; Score 25; DB 13; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 WTLK 10  
 DB 29 WTLK 32  
 RESULT 5  
 Q8VUW6  
 ID Q8VUW6 PRELIMINARY; PRT; 68 AA.  
 AC Q8VUW6  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE ORF13.  
 OS Staphylococcus hominis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Staphylococcus.  
 OK NCBI\_TaxID=1290;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC27844;



RA Katayama Y., Ito T., Hiramatsu K.;  
 RT "Localization of staphylococcal cassette chromosome in methicillin  
 RT susceptible *Staphylococci*.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB063171; BAB83492.1; 7842 MW; 578926E56D470DC CRC64;  
 SQ SEQUENCE 68 AA; 7842 MW; 578926E56D470DC CRC64;

Query Match 71.4%; Score 25; DB 2; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 65 WTLK 68

RESULT 6  
 Q9H1X5 PRELIMINARY; PRT; 70 AA.  
 ID Q9H1X5  
 AC Q9H1X5  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE BA16L21.1.3 (NADP-dependent leukotriene B4 12-hydroxydehydrogenase  
 DE (isoform 3)).  
 GN BA16L21.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laird G.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL135787; CAC22152.1; --  
 SQ SEQUENCE 70 AA; 7983 MW; E77FF1214CF13142 CRC64;

Query Match 71.4%; Score 25; DB 4; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

RESULT 7  
 Q8YWZ8 PRELIMINARY; PRT; 70 AA.  
 ID Q8YWZ8  
 AC Q8YWZ8  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Asl1421.  
 GN ASL1421.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213 (2001).  
 DR EMBL; AP003585; BAB73378.1; --  
 DR InterPro; IPR001387; HTH 3.  
 DR Pfam; PF01381; HTH 3; 1.  
 DR SMART; SM00530; HTH\_XRE; 1.

KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 70 AA; 8009 MW; 764CE5F9AC236DDA CRC64;

Query Match 71.4%; Score 25; DB 16; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 17 WTLK 20

RESULT 8  
 O98260 PRELIMINARY; PRT; 72 AA.  
 ID O98260  
 AC O98260;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE MHC class I antigen (Fragment).  
 GN GAMR-UA-PAL.4.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SPLEEN;  
 RX MEDLINE=20009526; PubMed=10541806;  
 RA Persson A.C., Stet R.J.M., Pilstrom L.;  
 RT "Characterization of MHC class I and beta2-microglobulin sequences in  
 RT Atlantic cod reveals an unusually high number of expressed class I  
 RT genes.";  
 RL Immunogenetics 50:49-59 (1999).  
 DR EMBL; AJ133450; CAB38040.1; --  
 DR InterPro; IPR001039; MHC I.  
 DR ProDom; PD000050; MHC\_I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 72 72  
 SQ SEQUENCE 72 AA; 8264 MW; FB260EAB403169BE CRC64;

Query Match 71.4%; Score 25; DB 7; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 69 WTLK 72

RESULT 9  
 Q919I3 PRELIMINARY; PRT; 89 AA.  
 ID Q919I3  
 AC Q919I3  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CUN094 hypothetical protein.  
 GN CUN094.  
 OS *Culex nigripalpus* baculovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
 OX NCBI\_TaxID=130556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FLORIDA1997;  
 RX MEDLINE=21488685; PubMed=11602755;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kutish G.F.;  
 RT "Genome Sequence of a Baculovirus Pathogenic for *Culex nigripalpus*.";  
 RL J. Virol. 75:11157-11165 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=FLORIDA1997;  
 RA Alfonso C.F., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kucish G.F.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF403738; AAK94172.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 89 AA; 10561 MW; 01F0A22D2C792541 CRC64;

Query Match 71.4%; Score 25; DB 12; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

Qy 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 10

Q9AR33 PRELIMINARY; PRT; 92 AA.  
 AC Q9AR33;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE P0537A05.26 Protein (P0554D10.2 protein).  
 GN P0537A05.26 OR P0554D10.2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone:P0537A05.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002869; BAB39218.1; -;  
 SQ SEQUENCE 92 AA; 10627 MW; 30D6BD50E14804DD CRC64;

Query Match 71.4%; Score 25; DB 10; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 24 WTLK 27

## RESULT 11

Q9HCV7 PRELIMINARY; PRT; 94 AA.  
 AC Q9HCV7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DJ1153D9.3 (novel protein).  
 GN DJ1153D9.3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Clark G.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL109806; CAC03742.1; -;  
 SQ SEQUENCE 94 AA; 11060 MW; E0AB7D817888A06C CRC64;

Query Match 71.4%; Score 25; DB 4; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 2 WTLK 5

## RESULT 12

Q67664 PRELIMINARY; PRT; 96 AA.  
 ID Q67664  
 AC Q67664;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF2 protein.  
 OS Garlic latent virus E29-6.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.  
 OX NCBI\_TaxID=143616;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Choi J.N., Choi Y.H., Choi Y.D., Lee J.S.;  
 RT "Nucleotide sequence of a cDNA for garlic mosaic virus.";  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X67134; CAA47614.1; -;  
 SQ SEQUENCE 96 AA; 10771 MW; A48BD84E2F7CF25A CRC64;

Query Match 71.4%; Score 25; DB 12; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

Qy 7 WTLK 10  
 ||||  
 Db 7 WTLK 10

## RESULT 13

Q91QC1 PRELIMINARY; PRT; 96 AA.  
 ID Q91QC1  
 AC Q91QC1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vpr protein.  
 GN VPR.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20338596; PubMed=10881687;  
 RA Yamada T., Iwamoto A.;

RT "Comparison of proviral accessory genes between long-term  
 nonprogressors and progressors of human immunodeficiency virus type 1  
 infection.";  
 RL Arch. Virol. 145:1021-1027(2000).  
 DR EMBL; AB034523; BAA93970.1; -;  
 DR InterPro; IPR000012; Retrov\_Vpr/X.  
 DR Pfam; PF00522; VPR; 1  
 DR PRINTS; PR00444; HIVPRVFX.  
 KW AIDS.

SQ SEQUENCE 96 AA; 11517 MW; 0C3C8C3A34FC8800 CRC64;

Query Match 71.4%; Score 25; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 14

Q8QDN8 PRELIMINARY; PRT; 96 AA.  
 AC Q8QDN8;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vpr protein.  
 GN VPR;  
 OS Simian-Human immunodeficiency virus.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=57667;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RX MEDLINE=21851104; PubMed=11861859;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RT "Evolution of a human immunodeficiency virus type 1 variant with  
 RT enhanced replication in pig-tailed macaque cells by DNA shuffling.";  
 RL J. Virol. 76:2924-2935(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1B3;  
 RA Pekrun K., Shibata R., Igarashi T., Reed M., Sheppard L., Patten P.A.,  
 RA Stemmer W.P.C., Martin M.A., Soong N.-W.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF465242; AAL78992.1; -;  
 SQ SEQUENCE 96 AA; 11451 MW; E41BB11BF66CDD6E CRC64;

Query Match 71.4%; Score 25; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 18 WTLK 21

## RESULT 15

Q85061 PRELIMINARY; PRT; 99 AA.  
 AC Q85061;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Papillomavirus sylvilagi Washington B (E2) and (E4) genes  
 DE (Fragment).  
 GN E4.  
 OS Cottontail rabbit papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RA Wu X., Brandsma J.;  
 RT "Papilloma Formation by Cottontail Rabbit Papillomavirus (CRPV)  
 RT Requires E1 and E2 Regulatory Genes in Addition to the E6 and E7  
 RT Transforming Genes.";  
 RL J. Virol. 0:0-0(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WASHINGTON B;  
 RX MEDLINE=85166175; PubMed=2984661;  
 RA Giri I., Danos O., Yaniv M.;  
 RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).

DR EMBL; U09496; AAB50542.2; -; V -> L (IN REF. 2).  
 FT CONFLICT 28 28  
 FT NON\_TER 99 99  
 SQ SEQUENCE 99 AA; 11346 MW; CE2A896101FB0BF0 CRC64;  
 Query Match 71.4%; Score 25; DB 12; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 73 WTLK 76

Search completed: January 29, 2003, 10:46:45  
 Job time : 23.4 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:37 ; Search time 9.45455 Seconds  
(without alignments)  
40.457 Million cell updates/sec

Title: US-09-707-738-23  
Perfect score: 34  
Sequence: 1 XXXXXWTLXXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgm2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgm2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	73.5	4	1	US-08-305-871A-17
2	25	73.5	4	4	US-08-788-822A-1
3	25	73.5	12	4	US-08-788-822A-1
4	25	73.5	13	1	US-08-305-871A-21
5	25	73.5	13	1	US-08-305-871A-22
6	25	73.5	13	4	US-08-788-822A-26
7	25	73.5	13	4	US-08-788-822A-27
8	25	73.5	14	4	US-08-788-822A-30
9	25	73.5	18	2	US-08-652-369A-1
10	25	73.5	23	4	US-07-741-453A-25
11	25	73.5	25	3	US-09-075-257A-17
12	25	73.5	25	3	US-09-534-639-17
13	25	73.5	28	3	US-09-075-257A-15
14	25	73.5	28	3	US-09-075-257A-16
15	25	73.5	28	4	US-09-534-639-15
16	25	73.5	28	4	US-09-534-639-16
17	25	73.5	86	4	US-08-858-200A-471
18	25	73.5	90	1	US-07-956-700B-85
19	25	73.5	90	1	US-08-476-537-85
20	25	73.5	90	1	US-08-485-607-85
21	25	73.5	90	2	US-08-475-879-85
22	25	73.5	90	4	US-09-433-043B-85
23	25	73.5	97	2	US-08-585-585A-7
24	25	73.5	97	2	US-08-249-037C-7
25	25	73.5	97	2	US-08-788-622B-7
26	25	73.5	97	3	US-08-788-621B-7
27	25	73.5	152	4	US-08-936-165A-264

Sequence 2948, Ap  
Sequence 3512, Ap  
Sequence 2944, Ap  
Sequence 54, Appl  
Sequence 376, Appl  
Sequence 374, Appl  
Sequence 58, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 2, Appl  
Sequence 17, Application US/08305871A  
Patent No. 5736142  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
APPLICANT: Gaeta, Federico  
APPLICANT: Grey, Howard M.  
APPLICANT: Sidney, John  
APPLICANT: Alexander, Jeffrey L.  
TITLE OF INVENTION: Alteration of Immune Response Using Pan  
TITLE OF INVENTION: DR-Binding Peptides  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,871A  
FILING DATE: 14-SEP-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/121,101  
FILING DATE: 14-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-0062-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-305-871A-17

## ALIGNMENTS

RESULT 1  
US-08-305-871A-17  
Sequence 17, Application US/08305871A  
Patent No. 5736142  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
APPLICANT: Gaeta, Federico  
APPLICANT: Grey, Howard M.  
APPLICANT: Sidney, John  
APPLICANT: Alexander, Jeffrey L.  
TITLE OF INVENTION: Alteration of Immune Response Using Pan  
TITLE OF INVENTION: DR-Binding Peptides  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,871A  
FILING DATE: 14-SEP-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/121,101  
FILING DATE: 14-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-0062-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-305-871A-17  
Query Match 73.5%; Score 25; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||||

Db 1 WTLK 4

RESULT 2  
US-08-788-822A-1  
; Sequence 1, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/788,822A  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010,510  
; FILING DATE: 24-JAN-1996  
; NAME: Bastian, Kevin L.  
; REGISTRATION/DOCKET NUMBER: 014137-009210US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010,510  
; FILING DATE: 24-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION/DOCKET NUMBER: 014137-009210US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-788-822A-1

Query Match 73.5%; Score 25; DB 4; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||||

Db 1 WTLK 4

RESULT 3  
US-08-788-822A-21  
; Sequence 21, Application US/08788822A  
; Patent No. 6413935  
; GENERAL INFORMATION:  
; APPLICANT: Alexander, Jeffrey L.  
; APPLICANT: DeFrees, Shawn  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Induction of Immune Response Against  
; TITLE OF INVENTION: Desired Determinants  
; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/788,822A  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010,510  
; FILING DATE: 24-JAN-1996  
; NAME: Bastian, Kevin L.  
; REGISTRATION/DOCKET NUMBER: 014137-009210US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 12  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "Xaa = alaninamide"  
US-08-788-822A-21

Query Match 73.5%; Score 25; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
|||||

Db 7 WTLK 10

RESULT 4  
US-08-305-871A-22  
; Sequence 22, Application US/08305871A  
; Patent No. 5736142  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Gaeta, Federico  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sidney, John  
; APPLICANT: Alexander, Jeffrey L.  
; TITLE OF INVENTION: Alteration of Immune Response Using Pan  
; TITLE OF INVENTION: DR-Binding Peptides  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA

```

;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305.871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; tyrosine or phenylalanine."
; US-08-305-871A-22
;
; Query Match 73.5%; Score 25; DB 1; Length 13;
; Best Local Similarity 100.0%; Pred. No. 33,
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 7 WTLK 10
; DB 7 WTLK 10
;
; RESULT 5
; US-08-305-871A-23
; Sequence 23, Application US/08305871A
; Patent No. 5736142
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Gaeta, Federico
; APPLICANT: Grey, Howard M.
; APPLICANT: Sidney, John
; APPLICANT: Alexander, Jeffrey L.
; TITLE OF INVENTION: Alteration of Immune Response Using Pan
; TITLE OF INVENTION: DR-Binding Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305.871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

;
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; tyrosine or phenylalanine."
; US-08-305-871A-23
;
; Query Match 73.5%; Score 25; DB 1; Length 13;
; Best Local Similarity 100.0%; Pred. No. 33,
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 7 WTLK 10
; DB 7 WTLK 10
;
; RESULT 6
; US-08-788-822A-26
; Sequence 26, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid

```





APPLICANT: Dean G. Tang, Kenneth V. Honn  
TITLE OF INVENTION: Induction of Tumor  
TITLE OF INVENTION: Cell Apoptosis With Chemical Inhibitors  
TITLE OF INVENTION: targeted to 12-lipoxygenase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ian C. McLeod  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 5.25 inch,  
MEDIUM TYPE: 360 kb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (version  
3.3)  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,369A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: Biomide 4.1-11  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE:  
DESCRIPTION: DNA primer for segment  
of 12-lipoxygenase.  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: human  
IMMEDIATE SOURCE:  
LIBRARY: Genomic  
US-08-652-369A-1

Query Match 73.5%; Score 25; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 1 WTLK 4

RESULT 10  
US-07-741-453A-25  
Sequence 25, Application US/07741453A  
Patent No. 6228597  
GENERAL INFORMATION:  
APPLICANT: PARMENTIER, MARC  
APPLICANT: LIBERT, FREDERIC  
APPLICANT: DUMONT, JACQUES  
APPLICANT: VASSART, GILBERT  
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR  
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS  
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: U.S.A.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/741,453A  
FILING DATE: 19911015  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-741-453A-25

Query Match 73.5%; Score 25; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 4 WTLK 7

RESULT 11  
US-09-075-257A-17  
Sequence 17, Application US/09075257A  
Patent No. 6074645  
GENERAL INFORMATION:  
APPLICANT: DIAMOND, DON JEFFREY  
APPLICANT: YORK, JOANNE  
TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BART G. NEWLAND  
STREET: 555 13TH STREET, NW SUITE 701E  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,257A  
FILING DATE: 11-MAY-1998  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/021,298  
FILING DATE: 10-FEB-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/950,064  
FILING DATE: 14-OCT-1997  
PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 08/747,488
; FILING DATE: 12-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NEWLAND, BART G
; REGISTRATION NUMBER: 31,282
; REFERENCE/DOCKET NUMBER: 1954-112CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 6
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine
; OTHER INFORMATION: or phenylalanine"
US-09-075-257A-17

Query Match 73.5%; Score 25; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 10 WTLK 13

RESULT 12
US-09-534-639-17
; Sequence 17, Application US/09534639
; Patent No. 6251399
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J
; APPLICANT: York, Joanne
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN
; TITLE OF INVENTION: CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-343
; CURRENT APPLICATION NUMBER: US/09/534,639
; CURRENT FILING DATE: 2000-03-27
; EARLIER APPLICATION NUMBER: 09/075,257
; EARLIER FILING DATE: 1998-05-11
; EARLIER APPLICATION NUMBER: 09/021,298
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: 08/950,064
; EARLIER FILING DATE: 1997-10-14
; EARLIER APPLICATION NUMBER: 08/747,488
; EARLIER FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (6)
; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human
; OTHER INFORMATION: Cytomegalovirus Vaccine
US-09-534-639-17

Query Match 73.5%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 40
```

```
Db 10 WTLK 13

RESULT 13
US-09-075-257A-15
; Sequence 15, Application US/09075257A
; Patent No. 6074645
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, DON JEFFREY
; APPLICANT: YORK, JOANNE
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BART G. NEWLAND
; STREET: 555 13TH STREET, NW SUITE 701E
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,257A
; FILING DATE: 11-MAY-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/021,298
; FILING DATE: 10-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/950,064
; FILING DATE: 14-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/747,488
; FILING DATE: 12-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NEWLAND, BART G
; REGISTRATION NUMBER: 31,282
; REFERENCE/DOCKET NUMBER: 1954-112CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 6
; OTHER INFORMATION: /note= "Xaa = cyclohexylalanine
; OTHER INFORMATION: or phenylalanine"
US-09-075-257A-15

Query Match 73.5%; Score 25; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10
Db 10 WTLK 13

RESULT 14
US-09-075-257A-16
; Sequence 16, Application US/09075257A
```

Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YORK, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G.  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 6  
; OTHER INFORMATION: /note= "xaa = cyclohexylalanine  
; OTHER INFORMATION: or phenylalanine"  
US-09-075-257A-16

Query Match 73.5%; Score 25; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 10 WTLK 13

RESULT 15  
US-09-534-639-15  
; Sequence 15, Application US/09534639  
; Patent No. 6251399  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J  
; APPLICANT: York, Joanne  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
; TITLE OF INVENTION: CYTOMEGALOVIRUS  
; FILE REFERENCE: 1954-343

; CURRENT APPLICATION NUMBER: US/09/534,639  
; CURRENT FILING DATE: 2000-03-27  
; EARLIER APPLICATION NUMBER: 09/075,257  
; EARLIER FILING DATE: 1998-05-11  
; EARLIER APPLICATION NUMBER: 09/021,298  
; EARLIER FILING DATE: 1998-02-10  
; EARLIER APPLICATION NUMBER: 08/950,064  
; EARLIER FILING DATE: 1997-10-14  
; EARLIER APPLICATION NUMBER: 08/747,488  
; EARLIER FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa = cyclohexylalanine or phenylalanine  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Human  
; OTHER INFORMATION: Cytomegalovirus Vaccine  
US-09-534-639-15

Query Match 73.5%; Score 25; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db 10 WTLK 13

Search completed: January 29, 2003, 10:48:22  
Job time : 10.4545 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:42:07 ; Search time 11.7091 Seconds  
(without alignments)  
114.943 Million cell updates/sec

Title: US-09-707-738-24  
Perfect score: 35  
Sequence: 1 XXXXXWTLKXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	71.4	14	2 A47421	leukotriene B-4 12
2	25	71.4	70	2 AB1984	hypothetical prote
3	25	71.4	81	2 S3586	high potential iro
4	25	71.4	85	1 IHKREV	high potential iro
5	25	71.4	85	1 IHTFER	high potential iro
6	25	71.4	90	2 D97355	hypothetical prote
7	25	71.4	96	2 S3863	hypothetical prote
8	25	71.4	104	2 A41384	nicotinic acetylch
9	25	71.4	122	2 A45385	translation repres
10	25	71.4	127	2 C97299	hypothetical prote
11	25	71.4	134	2 T12858	hypothetical prote
12	25	71.4	137	2 G90472	partial transposas
13	25	71.4	152	2 F90143	hypothetical prote
14	25	71.4	152	2 C90295	hypothetical prote
15	25	71.4	159	2 E72731	hypothetical prote
16	25	71.4	160	2 T23232	hypothetical prote
17	25	71.4	163	2 S86658	hypothetical prote
18	25	71.4	164	2 H89784	hypothetical prote
19	25	71.4	172	2 G99764	conserved hypotet
20	25	71.4	173	2 A83965	hypothetical prote
21	25	71.4	175	2 T48641	transposase (21) B
22	25	71.4	177	2 T00789	hypothetical prote
23	25	71.4	180	2 T46695	ubiquitin-protein
24	25	71.4	187	2 E69170	hypothetical prote
25	25	71.4	201	2 F97063	hypothetical prote
26	25	71.4	201	2 AD3302	surfactin biosynth
27	25	71.4	201	2 C97558	31k outer-membrane
28	25	71.4	201	2 AG2778	hypothetical prote
29	25	71.4	203	2 C69063	conserved hypotet

30 25 71.4 207 2 C70029 hypothetical prote  
31 25 71.4 211 2 H69539 SSU ribosomal prot  
32 25 71.4 212 1 W4WLRB E4 protein - cotto  
33 25 71.4 217 2 AF1898 hypothetical prote  
34 25 71.4 218 2 D71693 hypothetical prote  
35 25 71.4 219 2 AC3615 31k outer-membrane  
36 25 71.4 238 2 I67638 proteasome activat  
37 25 71.4 239 2 I53518 proteasome activat  
38 25 71.4 240 2 T45814 hypothetical prote  
39 25 71.4 243 2 T23077 hypothetical prote  
40 25 71.4 252 2 D87397 hypothetical prote  
41 25 71.4 263 2 S06330 ricin E - castor b  
42 25 71.4 264 2 B95259 conserved hypotet  
43 25 71.4 264 2 E98124 hypothetical prote  
44 25 71.4 266 2 H98208 probable permease  
45 25 71.4 266 2 AI3077 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A47421  
leukotriene B-4 12-hydroxydehydrogenase (EC 1.1.1.1.-) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 19-May-1995 #sequence\_revision 19-May-1995 #text\_change 17-Mar-1999  
C:Accession: A47421  
R:Yokomizo, T.; Izumi, T.; Takahashi, T.; Kasama, T.; Kobayashi, Y.; Sato, F.; Taketani, J. Biol. Chem. 268, 18128-18135, 1993  
A:Title: Enzymatic inactivation of leukotriene B-4 by a novel enzyme found in the porcine  
A:Reference number: A47421; MUID:93352633; PMID:8394361  
A:Accession: A47421  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <YOK>  
C:Keywords: oxidoreductase

Query Match 71.4%; Score 25; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db |||||  
6 WTLK 9

##### RESULT 2

AB1984  
hypothetical protein asl1421 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AB1984  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AB1984  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-70 <DKR>  
A:Cross-references: DB:BA000019; PIDN:BA073378.1; PID:gl7130768; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: asl1421

Query Match 71.4%; Score 25; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
Db |||||

Db 17 WTLK 20

RESULT 3

high potential iron-sulfur protein - Chromatium tepidum

C:Date: 10-Dec-1993 #sequence\_revision 27-feb-1997 #text\_change 18-Jun-1999

C:Accession: S35586

R:Mouill, J.M.; Scherrer, N.; Gagnon, J.; Forest, E.; Pettillot, Y.; Garcia, D.

Arch. Biochem. Biophys. 305, 186-192, 1993

A:Title: Primary structure of Chromatium tepidum high-potential iron-sulfur protein in

A:Reference number: S35586; MUID:93343628; PMID:8393645

A:Accession: S35586

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-81 <MOU>

C:Superfamily: high potential iron-sulfur protein; metalloprotein

C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein

F:43,46,61,75/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 71.4%; Score 25; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 78 WTLK 81

RESULT 4

THKREV

high potential iron-sulfur protein [validated] - Chromatium vinosum

N:Alternate names: HiPIP

C:Species: Chromatium vinosum

C:Date: 24-Apr-1984 #sequence\_revision 24-Oct-1997 #text\_change 15-Sep-2000

C:Accession: A92330; A00263

R:Tedro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.

J. Biol. Chem. 256, 731-735, 1981

A:Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur

A:Reference number: A92330; MUID:81094036; PMID:7451471

A:Accession: A92330

A:Molecule type: protein

A:Residues: 1-73, 'D', 75-85 <TED>

R:Due, K.; Tedro, S.; Bartsch, R.G.

J. Biol. Chem. 248, 7318-7331, 1973

A:Title: The complete amino acid sequence of Chromatium high potential iron sulfur prote

A:Reference number: A92143; MUID:74012043; PMID:4745771

A:Accession: A92143

A:Molecule type: protein

A:Residues: 1-10, 'N', 12-44, 'D', 46-85 <DUS>

A:Experimental source: strain D

R:Banci, L.; Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Somporpnisut, P.

submitted to the Brookhaven Protein Data Bank, January 1995

A:Reference number: A65814; PDB:1HRQ

A:Contents: annotation; conformation by (1)H-NMR, reduced form, residues 1-85

R:Bertini, I.; Dikiy, A.; Kastrau, D.H.W.; Luchinat, C.; Somporpnisut, P.

submitted to the Brookhaven Protein Data Bank, December 1995

A:Reference number: A66207; PDB:1NEH

A:Contents: annotation; conformation by (1)H-NMR, oxidized form, residues 1-85

R:Backes, G.; Mino, Y.; Loehr, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman

J. Am. Chem. Soc. 113, 2055-2064, 1991

A:Title: The environment of Fe4S4 clusters in ferredoxins and high-potential iron protei

A:Reference number: A44688

A:Contents: annotation; X-ray crystallography, 2.0 angstroms

A:Note: assignment of Raman spectra frequencies and hydrogen bonds around the iron-sulfu

R:Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Xuong, N.H.; Alden, R.A.; Bartsch, R.G.

J. Biol. Chem. 249, 4212-4225, 1974

A:Title: Two-angstrom crystal structure of oxidized Chromatium high potential iron prote

A:Reference number: A92153; MUID:74309824; PMID:4855287

A:Contents: annotation; X-ray crystallography, 2.0 angstroms

R:Carter Jr., C.W.; Kraut, J.; Freer, S.T.; Alden, R.A.

J. Biol. Chem. 249, 6339-6346, 1974

A:Title: Comparison of oxidation-reduction site geometries in oxidized and reduced Chrome

A:Reference number: A92161; MUID:75019502; PMID:4417854

A:Contents: annotation; X-ray crystallography

A:Note: structures of the oxidized and reduced forms are compared with each other and wit

C:Superfamily: high potential iron-sulfur protein; metalloprotein

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F:43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental

Query Match 71.4%; Score 25; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 80 WTLK 83

RESULT 5

ITHFER

high potential iron-sulfur protein - Thiocapsa roseopersicina (tentative sequence)

C:Species: Thiocapsa roseopersicina

C:Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 31-Mar-2000

C:Accession: A00264

R:Tedro, S.M.; Meyer, T.E.; Bartsch, R.G.; Kamen, M.D.

J. Biol. Chem. 256, 731-735, 1981

A:Title: Primary structures of high potential, four-iron-sulfur ferredoxins from the pur

A:Reference number: A92330; MUID:81094036; PMID:7451471

A:Accession: A00264

A:Molecule type: protein

A:Residues: 1-85 <TED>

C:Superfamily: high potential iron-sulfur protein; metalloprotein

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F:43,46,63,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 71.4%; Score 25; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 80 WTLK 83

RESULT 6

D97355

hypothetical protein CAC3711 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: D97355

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D97355

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <KUP>

A:Cross-references: GB:AE001437; PIDN:AAK81631.1; PID:G15026816; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3711

Query Match 71.4%; Score 25; DB 2; Length 90;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 46 WTLK 49

RESULT 7

## S23863

hypothetical protein 2 - garlic mosaic virus  
 C:Species: Garlic mosaic virus, GarMV  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
 C:Accession: S23863  
 R:Choi, J.N.; Choi, Y.H.; Choi, Y.D.; Lee, J.S.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Nucleotide sequence of a cDNA for garlic mosaic virus.  
 A:Reference number: S23862  
 A:Accession: S23863  
 A>Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 1-96 <CHO>  
 A:Cross-references: EMBL:X67134; NID:g59391; PIDN:CAA47614.1; PID:g59393

Query Match 71.4%; Score 25; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 7 WTLK 10

## RESULT 8

A41384  
 nicotinic acetylcholine receptor alpha chain - Chinese cobra (fragment)  
 C:Species: Naja naja atra (Chinese cobra)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: A41384  
 R:Neumann, D.; Barchan, D.; Horowitz, M.; Kochva, E.; Fuchs, S.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7255-7259, 1989  
 A:Title: Snake acetylcholine receptor: cloning of the domain containing the four extracellular loops  
 A:Reference number: A41384; MUID:89386720; PMID:2780569  
 A:Accession: A41384  
 A:Molecule type: mRNA  
 A:Residues: 1-104 <NEU>  
 A:Cross-references: GB:M26388; NID:g213370; PIDN:AAA49384.1; PID:g213371  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: neurotransmitter receptor

Query Match 71.4%; Score 25; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 58 WTLK 61

## RESULT 9

A45385  
 translation repressor regA - phage RB69  
 C:Species: phage RB69  
 C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: A45385  
 R:Jozwik, C.E.; Miller, E.S.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 5053-5057, 1992  
 A:Title: Regions of bacteriophage T4 and RB69 regA translational repressor proteins that interact with the 3' noncoding region of the T4 genome  
 A:Reference number: A45385; MUID:92279264; PMID:1594613  
 A:Accession: A45385  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-122 <JOZ>  
 A:Cross-references: GB:M86231; NID:g215354; PIDN:AAA32295.1; PID:g215356  
 C:Genetics:  
 A:Gene: regA  
 C:Superfamily: phage T4 translation repressor  
 C:Keywords: RNA binding; translation repressor

Query Match 71.4%; Score 25; DB 2; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 112 WTLK 115

## RESULT 10

C97299  
 hypothetical protein CAC3248 [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: C97299  
 R:Nolling, J.; Bennett, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: C97299  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-127 <KUR>  
 A:Cross-references: GB:AB001437; PIDN:AAK81182.1; PID:g15026321; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3248

Query Match 71.4%; Score 25; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10

Db 103 WTLK 106

## RESULT 11

T12858  
 hypothetical protein yopX - Bacillus subtilis phage SPBc2  
 C:Species: Bacillus subtilis phage SPBc2  
 C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: T12858; F69918  
 R:Lazarevic, V.; Dueterhoeft, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D.  
 submitted to the EMBL Data Library, August 1997  
 A:Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 prophage  
 A:Reference number: 217583  
 A:Accession: T12858  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-134 <LAZ>

A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025478; PID:AAI3067.1  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, A.; Authors: Schleicher, S.; Schroeter, R.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaumoto, K.; Yata, K.; Yoshida, K.  
 A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: F69918  
 A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-134 &lt;KUN&gt;

A:Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB13991.1; PH  
 A:Experimental source: strain 168  
 C:Genetics:

A:Gene: yopX  
C:Superfamily: Bacillus subtilis phage SPBc2 hypothetical protein yopX

Query Match 71.4%; Score 25; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 31 WTLK 34

RESULT 12  
G90472  
partial transposase ISC1078 [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: G90472  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: G90472  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-137 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3816304; PIDN:AAK43038.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO2933

Query Match 71.4%; Score 25; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

RESULT 13  
F90143  
hypothetical protein SSO0040 [imported] - Sulfolobus solfataricus transposon ISC1078  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 18-Jul-2001  
C:Accession: F90143; G90336; G90480  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: F90143  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3813171; PIDN:AAK40405.1; GSPDB:GN00155  
A:Accession: G90336  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <K2>  
A:Cross-references: GB:AE006641; NID:gl3814994; PIDN:AAK41946.1; GSPDB:GN00155  
A:Accession: G90480  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <K3>  
A:Cross-references: GB:AE006641; NID:gl3816382; PIDN:AAK43098.1; GSPDB:GN00155  
A:Accession: G90280  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <K4>  
A:Cross-references: GB:AE006641; NID:gl3814460; PIDN:AAK41502.1; GSPDB:GN00155

C:Genetics:  
A:Gene: SSO0040; SSO1750; SSO2992; SSO1264

Query Match 71.4%; Score 25; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

RESULT 14  
C90295  
hypothetical protein SSO1384 [imported] - Sulfolobus solfataricus transposon ISC1078  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: C90295  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: C90295  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <KUR>  
A:Cross-references: GB:AE006641; NID:gl3814600; PIDN:AAK41618.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO1384

Query Match 71.4%; Score 25; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 91 WTLK 94

RESULT 15  
E72731  
hypothetical protein APE0390 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: E72731  
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha-  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku-  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyru  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: E72731  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-159 <KAW>  
A:Cross-references: DDBJ:AP000059; NID:gs103911; PIDN:BAA79345.1; PID:gs104029  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0390

Query Match 71.4%; Score 25; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
|||  
Db 37 WTLK 40

Search completed: January 29, 2003, 10:47:36  
Job time : 11.7091 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 10:30:17 ; Search time 6.36364 Seconds  
(without alignments)  
91.248 Million cell updates/sec

Title: US-09-707-738-24

Perfect score: 35

Sequence: 1 XXXXXXXWTLKXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	71.4	83	1 HPIS_THETI	P80176 thermochrom
2	25	71.4	85	1 HPIS_THIRO	P00261 thiocapsa r
3	25	71.4	90	1 YIBI_CLOAB	P04353 clostridium
4	25	71.4	98	1 NULM_BRARE	Q4my2 brachydanio
5	25	71.4	104	1 ACPS_NAJNA	P14143 najja naja (
6	25	71.4	114	1 ACPS_MYCGE	Q9zb79 mycoplasma
7	25	71.4	122	1 HPIS_CHRVI	P00260 chromatium
8	25	71.4	122	1 REGA_BPR69	Q01751 bacterioph
9	25	71.4	125	1 CP03_HUMAN	O95177 homo sapien
10	25	71.4	211	1 CP03_ARCFU	O27964 archaeoglob
11	25	71.4	212	1 V84_CRPVK	P03124 cottoncail
12	25	71.4	218	1 Y363_RICPR	Q9zd96 rickettsia
13	25	71.4	233	1 MSRA_BOVIN	P54149 bos taurus
14	25	71.4	238	1 PSE2_RAT	Q63798 rattus norv
15	25	71.4	239	1 PSE2_HUMAN	Q9ul46 homo sapien
16	25	71.4	239	1 PSE2_MOUSE	P97372 mus musculu
17	25	71.4	240	1 OM31_BRUME	O45322 brucella me
18	25	71.4	282	1 YC80_GUITH	Q78449 guillardia
19	25	71.4	294	1 UP08_ECOLI	P39173 escherichia
20	25	71.4	301	1 LEC1_RAT	P08290 rattus norv
21	25	71.4	309	1 YC30_ODOSI	P49518 odontella s
22	25	71.4	311	1 LB4D_HUMAN	Q14914 homo sapien
23	25	71.4	329	1 LB4D_PIG	Q29073 sus scrofa
24	25	71.4	334	1 YEDY_ECO57	O8xb74 escherichia
25	25	71.4	334	1 YEDY_ECOLI	P76342 escherichia
26	25	71.4	334	1 YEDY_SALTY	Q9xsi salmonella
27	25	71.4	345	1 ARGC_BACHD	Q9K8v2 bacillus ha
28	25	71.4	346	1 CATV_GVXN	Q9py75 xestia c-ni
29	25	71.4	349	1 I105_MOUSE	P61190 mus musculu
30	25	71.4	349	1 LB4D_RABIT	Q28719 oryctolagus
31	25	71.4	349	1 MRAY_CHLPN	Q9z706 chlamydia p
32	25	71.4	349	1 Y028_BORBU	O51059 borrelia bu
33	25	71.4	352	1 THBC_HUMAN	P55103 homo sapien

RESULT 1					
HPIS_THETI	34	25	71.4	353	1 PSBD_MAIZE
ID	35	25	71.4	355	1 VAL1_ABMVW
AC	36	25	71.4	357	1 O2B2_HUMAN
P80176;	37	25	71.4	360	1 VE2_HPV70
DT 01-OCT-1993 (Rel. 27, Created)	38	25	71.4	361	1 VAL1_PYMVW
DT 01-OCT-1993 (Rel. 27, Last sequence update)	39	25	71.4	365	1 HHP1_SCHPO
DT 15-JUN-2002 (Rel. 41, Last annotation update)	40	25	71.4	370	1 VE2_HPV39
DE High potential iron-sulfur protein (HiPIP).	41	25	71.4	372	1 3BHS_MACMU
GN HiP.	42	25	71.4	372	1 ARIB_HUMAN
OS Thermochromatium tepidum (Chromatium tepidum).	43	25	71.4	372	1 ARIB_MOUSE
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;	44	25	71.4	372	1 ARIB_RAT
OC Thermochromatium.	45	25	71.4	373	1 NOLL_RHILO
OX NCBI_TaxID=1050;					
RN [1]					
RP SEQUENCE.					
RC STRAIN=ATCC 43061;					
RX MEDLINE=93343628; PubMed=8933645;					
RA Moulis J.-M., Scherrer N., Gagnon J., Forest E., Petillot Y.,					
RA Garcia D.;					
RT "Primary structure of Chromatium tepidum high-potential iron-sulfur					
RT protein in relation to thermal denaturation.";					
RL Arch. Biochem. Biophys. 305:186-192(1993).					
RN [2]					
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).					
RX MEDLINE=20558546; PubMed=1095707;					
RA Nogi T., Fatmir I., Kobayashi M., Nozawa T., Miki K.;					
RT "Crystal structures of photosynthetic reaction center and high-					
RT potential iron-sulfur protein from Thermochromatium tepidum:					
RT thermostability and electron transfer.";					
Proc. Natl. Acad. Sci. U.S.A. 97:13561-13566(2000).					
RN [3]					
RP X-RAY CRYSTALLOGRAPHY (0.8 ANGSTROMS).					
RX MEDLINE=22071404; PubMed=12077426;					
RA Liu L., Nogi T., Kobayashi M., Nozawa T., Miki K.;					
RT "Ultra-high-resolution structure of high-potential iron-sulfur					
RT protein from Thermochromatium tepidum.";					
Acta Crystallogr. D 58:1085-1091(2002).					
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S					
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST					
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST					
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.					
CC -!- SUBUNIT: HOMODIMER (PROBABLE).					
CC -!- MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +323 MV.					
CC -!- MISCELLANEOUS: THIS PROTEIN IS THERMOSTABLE.					
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN					
(HIPIP) FAMILY.					
CC PIR; S35586; S35586.					
DR PDB; 1EYT; 13-DEC-00.					
DR PDB; 1LUA; 20-MAR-02.					
DR InterPro; IPR000170; Hipot_iron_sulf.					
DR Pfam; PF01355; HIPIP; 1.					
DR PRINTS; PR00374; HIPIPFDOXIN.					
DR PROSITE; PS00596; HIPIP; 1.					
KW Electron transport; iron-sulfur; 4Fe-4S; 3D-structure.					
FT METAL 43 43					
FT METAL 46 46					
FT IRON-SULFUR (4FE-4S).					
FT IRON-SULFUR (4FE-4S).					

#### ALIGNMENTS

P48184 zea mays (m  
P21947 abutilon mo  
Q9gk3 homo sapien  
P50773 human papil  
P27258 potato yell  
P40235 schizosacch  
P24830 human papil  
P27365 m 3 beta-hy  
O15143 homo sapien  
Q9w32 mus musculu  
O8656 rattus norv  
Q52778 rhizobium 1

FT METAL 61 61 IRON-SULFUR (4FE-4S).  
 FT METAL 75 75 IRON-SULFUR (4FE-4S).  
 SQ SEQUENCE 83 AA; 8786 MW; 9211684FD2C44E0A CRC64;

Query Match 71.4%; Score 25; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 78 WTLK 81

## RESULT 2

ID HPIS THIRO STANDARD; PRT; 85 AA.  
 AC P00261;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein (HiPIP).  
 GN HiP.  
 OS Thiocapsa roseopersicina.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae; Thiocapsa.  
 OX NCBI\_TaxID=1058;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=81094036; PubMed=7451471;  
 RA Tedro S.M., Meyer T.E., Bartech R.G., Kamen M.D.;  
 RT "Primary structures of high potential, four-iron-sulfur ferredoxins  
 RT from the purple sulfur photosynthetic bacteria, Thiocapsa  
 RT roseopersicina and Chromatium gracile.";  
 RL J. Biol. Chem. 256:731-735(1981).  
 CC -1- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
 CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 CC (HIPIP) FAMILY.  
 DR PIR; A00264; IHTPER.  
 DR HSPP; P00260; ICKU.  
 DR InterPro; IPR000170; Hipot\_iron\_sulf.  
 DR Pfam; PF01355; HiPIP; 1.  
 DR PRINTS; PR00374; HIPIPRDOXIN.  
 DR PROSITE; PS00596; HiPIP; 1.

FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 FT METAL 46 46 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 FT METAL 63 63 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 FT METAL 77 77 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 SQ SEQUENCE 85 AA; 8889 MW; C47AD747D2218482 CRC64;  
 Query Match 71.4%; Score 25; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 80 WTLK 83

## RESULT 3

Y1B1 CLOAB STANDARD; PRT; 90 AA.  
 ID Y1B1 CLOAB  
 AC Q04353;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein CAC3711.  
 GN CAC3711.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=93273706; PubMed=8501044;  
 RA Sauer U., Duerre P.;  
 RT "Sequence and molecular characterization of a DNA region encoding a  
 RT small heat shock protein of Clostridium acetobutylicum.";  
 RL J. Bacteriol. 175:3394-3400(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466286;  
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium Clostridium acetobutylicum.";  
 RL J. Bacteriol. 183:4823-4838(2001).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X65276; CAA46377.1; ALT INIT.  
 DR EMBL; AE007866; AAK81631.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 90 AA; 10302 MW; DDDDE0F5C2PDD3CC CRC64;

Query Match 71.4%; Score 25; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WTLK 10  
 ||||  
 Db 46 WTLK 49

## RESULT 4

ID NULM BRARE STANDARD; PRT; 98 AA.  
 AC Q9MIY2;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
 GN MTND4L OR ND4L.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB;  
 RX MEDLINE=21549095; PubMed=11691861;  
 RA Broughton R.E., Milam J.E., Roe B.A.;  
 RT "The complete sequence of the zebrafish (Danio rerio) mitochondrial  
 RT genome and evolutionary patterns in vertebrate mitochondrial DNA.";  
 RL Genome Res. 11:1958-1967(2001).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AC024175; AAF74305.1; -;  
 DR ZFIN; ZDB-GENE-011205-11; mtdn4L.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidored4L.  
 DR Pfam; PF00420; oxidored\_q2; 1.  
 DR ProDom; PD000359; Oxidored4L; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 98 AA; 10523 MW; 54C4B6230D23EF24 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 WTLK 10  
 ||||  
 Db 47 WTLK 50

## RESULT 5

ACHA\_NAJNA  
 ID ACHA\_NAJNA STANDARD; PRT; 104 AA.  
 AC P14143;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acetylcholine receptor protein, alpha chain (Fragment).  
 OS Naja naja (Indian cobra).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroides;  
 OC Elapidae; Elapinae; Naja.  
 OX NCBI\_TaxID=35670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9396720; PubMed=2780569;  
 RA Neumann B., Barchan B., Horowitz M., Kochva E., Fuchs S.;  
 RT "Snake acetylcholine receptor: cloning of the domain containing the  
 RT four extracellular cysteines of the alpha subunit".;  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7255-7259(1989).  
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA CHAINS.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M26388; AAA49384.1; -;  
 DR PIR; A41384; A41384.  
 DR InterPro; IPR001175; Neur channel.  
 DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
 DR PROSITE; PS00236; NEUOTR\_ION\_CHANNEL; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein.  
 FT NON TER 1  
 FT DOMAIN <1>104 EXTRACELLULAR.  
 FT DISULFID 10 24 BY SIMILARITY.  
 FT DISULFID 74 75 ASSOCIATED WITH RECEPTOR ACTIVATION.  
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT NON TER 104 104  
 SQ SEQUENCE 104 AA; 12194 MW; 47A39E2C9BFA7A0 CRC64;

Oy 7 WTLK 10  
 ||||  
 Db 58 WTLK 61

## RESULT 6

ACPS MYCGE  
 ID ACPS MYCGE STANDARD; PRT; 114 AA.  
 AC Q9ZB79;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Holo-[acyl-carrier protein] synthase (EC 2.7.8.7) (Holo-ACP synthase)  
 DE (4'-phosphopantetheinyl transferase acps).  
 GN ACPS OR MG211.1  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Frichman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium".;  
 RL Science 270:397-403(1995).  
 RN [2]  
 RP IDENTIFICATION.

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Frichman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Transfers the 4'-phosphopantetheine moiety from coenzyme  
 CC A to a Ser of acyl-carrier protein (By similarity).  
 CC -!- CATALYTIC ACTIVITY: CoA + apo-[acyl-carrier protein] = adenosine  
 CC 3',5'-bisphosphate + holo-[acyl-carrier protein].  
 CC -!- COFACTOR: Magnesium (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY. ACPS  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U39701; AAC71438.1; -;  
 DR TIGR; MG211.1; -;  
 DR InterPro; IPR002582; ACPS.  
 DR InterPro; IPR004568; Pantethn\_trn.  
 DR Pfam; PF01648; ACPS; 1.  
 DR ProDom; PD004282; ACPS; 1.  
 DR TIGRFAMs; TIGR00556; pantethn\_trn; 1.  
 KW Transferee; Lipid synthesis; Fatty acid biosynthesis; Magnesium;  
 FT METAL 8 8 MAGNESIUM (BY SIMILARITY).  
 FT METAL 58 58 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 114 AA; 13153 MW; 8992BBB560373B92 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 |||||  
 DB 54 WTLK 57

## RESULT 7

ID HTIS CHRVI STANDARD; PRT; 122 AA.  
 AC P00260; P96753; Q9R4K3;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE High potential iron-sulfur protein precursor (HiPIP).  
 GN HiP.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochromatium.  
 CX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=97320627; PubMed=9177478;  
 RA Bruser T., Truper H.G., Dahl C.;  
 RT "Cloning and sequencing of the gene encoding the high potential iron-  
 sulfur protein (HiPIP) from the purple sulfur bacterium Chromatium  
 vinosum.";  
 RL Biochim. Biophys. Acta 1352:18-22 (1997).  
 RN [2]  
 RP SEQUENCE OF 38-122.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=74012043; PubMed=4745771;  
 RA Dus K., Tedro S., Bartsch R.G.;  
 RT "The complete amino acid sequence of Chromatium high potential iron  
 sulfur protein.";  
 RL J. Biol. Chem. 248:7318-7331 (1973).  
 RN [3]  
 RP REVISIONS TO 48; 82 AND 111.  
 RX MEDLINE=81094036; PubMed=7451471;  
 RA Tedro S.M., Meyer T.E., Bartsch R.G., Kamen M.D.;  
 RT "Primary structures of high potential, four-iron-sulfur ferredoxins  
 from the purple sulfur photosynthetic bacteria, Thiocapsa  
 roseopersicina and Chromatium gracile.";  
 RL J. Biol. Chem. 256:731-735 (1981).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=74309824; PubMed=4855287;  
 RA Carter C.W. Jr., Kraut J., Freer S.T., Xuong N.H., Alden R.A.,  
 RA Bartsch R.G.;  
 RT "2-A crystal structure of oxidized Chromatium high potential iron  
 protein.";  
 RL J. Biol. Chem. 249:4212-4225 (1974).  
 RN [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92135210; PubMed=1734968;  
 RA Nettesheim D.G., Harder S.R., Feinberg B.A., Otvos J.D.;  
 RT "Sequential resonance assignments of oxidized high-potential  
 iron-sulfur protein from Chromatium vinosum.";  
 RL Biochemistry 31:1234-1244 (1992).  
 RN [6]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=92304939; PubMed=1610810;  
 RA Gallard J., Albrand J.-P., Moullis J.-M., Wemmer D.E.;  
 RT "Sequence-specific assignments of the 1H nuclear magnetic resonance  
 spectra of reduced high-potential ferredoxin (HiPIP) from Chromatium  
 vinosum.";  
 RL Biochemistry 31:5632-5639 (1992).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95118989; PubMed=7819198;

RA Banci L., Bertini I., Dikiy A., Kastrau D.H.W., Luchinat C.,  
 RA Sompornpisut P.;  
 RT "The three-dimensional solution structure of the reduced high-  
 potential iron-sulfur protein from Chromatium vinosum through NMR.";  
 RL Biochemistry 34:206-219 (1995).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=95359150; PubMed=7632685;  
 RA Bertini I., Dikiy A., Kastrau D.H., Luchinat C., Sompornpisut P.;  
 RT "Three-dimensional solution structure of the oxidized high potential  
 iron-sulfur protein from Chromatium vinosum through NMR. Comparative  
 analysis with the solution structure of the reduced species.";  
 RL Biochemistry 34:9851-9858 (1995).  
 RN [9]  
 RP STRUCTURE BY NMR OF MUTANT SER-114.  
 RX MEDLINE=96216867; PubMed=8639555;  
 RA Bentrup D., Bertini I., Cappozzi F., Dikiy A., Eltis L., Luchinat C.;  
 RT "Three-dimensional structure of the reduced C77S mutant of the  
 Chromatium vinosum high-potential iron-sulfur protein through nuclear  
 magnetic resonance: comparison with the solution structure of the  
 wild-type protein.";  
 RL Biochemistry 35:5928-5936 (1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (0.93 ANGSTROMS).  
 RX MEDLINE=20003126; PubMed=10531472;  
 RA Parisini E., Capozzi F., Lubini F., Lanzin V., Luchinat C.,  
 RA Sheldrick G.M.;  
 RT "Ab initio solution and refinement of two high-potential iron protein  
 structures at atomic resolution.";  
 RL Acta Crystallogr. D 55:1773-1784 (1999).  
 CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S  
 FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST  
 PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST  
 ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.  
 CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +360 MV.  
 CC -!- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN  
 (HIPIP) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U01381; AAB48829.1; -.  
 DR PIR; A00263; IHKREV.  
 DR PDB; IHIP; 15-APR-93.  
 DR PDB; IHRQ; 03-JUN-95.  
 DR PDB; IHRH; 31-JUL-95.  
 DR PDB; INEH; 08-MAR-96.  
 DR PDB; INOE; 10-JUN-96.  
 DR PDB; INOE; 16-DEC-98.  
 DR PDB; ICKU; 13-MAY-99.  
 DR InterPro; IPR000170; Hipot\_ironsulph.  
 DR Pfam; PF01385; HIPIP; 1.  
 DR PRINTS; PR00374; HIPIPFRRDOXIN.  
 DR PROSITE; PS00596; HIPIP; 1.  
 KW Electron transport; Iron-sulfur; 4Fe-4S; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 37  
 FT CHAIN 38 122 HIGH POTENTIAL IRON-SULFUR PROTEIN.  
 FT METAL 80 80 IRON-SULFUR (4FE-4S).  
 FT METAL 83 83 IRON-SULFUR (4FE-4S).  
 FT METAL 100 100 IRON-SULFUR (4FE-4S).  
 FT METAL 114 114 IRON-SULFUR (4FE-4S).  
 FT CONFLICT 111 111 N -> D (IN REF. 3).  
 FT TURN 41 42  
 FT TURN 43 43  
 FT STRAND 46 47  
 FT TURN 46 47

FT HELIX 49 54  
 FT TURN 55 55  
 FT STRAND 57 57  
 FT HELIX 60 62  
 FT HELIX 65 68  
 FT HELIX 75 77  
 FT HELIX 80 82  
 FT STRAND 83 83  
 FT TURN 84 85  
 FT STRAND 86 87  
 FT TURN 88 89  
 FT TURN 91 92  
 FT STRAND 97 100  
 FT TURN 101 102  
 FT TURN 104 105  
 FT STRAND 107 109  
 FT TURN 110 111  
 FT STRAND 113 113  
 FT TURN 115 116  
 FT STRAND 119 119  
 SQ SEQUENCE 122 AA; 12761 MW; ACBBAFC917F32A09 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 117 WTLK 120

## RESULT 8

ID REGA\_BPR69 STANDARD; PRT; 122 AA.  
 AC Q01751;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Translation repressor protein.  
 GN REGA.  
 OS Bacteriophage RB69.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 OC T4-like viruses.  
 OX NCBI\_TaxID=12353;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=92279264; PubMed=1594613;  
 RA Jozwik C.E., Miller E.S.;  
 RT "Regions of bacteriophage T4 and RB69 RegA translational repressor  
 proteins that determine RNA-binding specificity.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5053-5057(1992).  
 CC -!- FUNCTION: CONTROLS THE TRANSLATION OF A NUMBER OF PROTEINS (SUCH  
 AS REGA ITSELF, RIIB AND AT LEAST 35 OTHERS) BY BINDING TO THEIR  
 MRNA.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M86231; AAA32295.1; -;  
 DR HSSP; P04528; 1REG.  
 DR InterPro; IPR002702; Translat\_reg.  
 DR Pfam; PF01818; Translat\_reg; 1.  
 DR ProDom; PD031547; Translat\_reg; 1.  
 KW Translation regulation; Repressor; DNA-binding.  
 FT DNA BIND 15 37 H-T-H MOTIF (POTENTIAL).  
 FT MUTAGEN 24 24 I->T: ACTIVITY ALTERED.  
 FT MUTAGEN 25 25 A->V: 100% ACTIVITY LOSS.  
 FT MUTAGEN 37 37 H->Y: 100% ACTIVITY LOSS.

FT MUTAGEN 72 72 D->G: 100% ACTIVITY LOSS.  
 SQ SEQUENCE 122 AA; 14431 MW; C4374CE90E16D1D3 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 112 WTLK 115

## RESULT 9

CP03\_HUMAN STANDARD; PRT; 125 AA.  
 ID CP03\_HUMAN  
 AC O95177; O95176;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protein Cl6orf3.  
 DE Cl6ORF3.  
 GN Cl6ORF3.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99009317; PubMed=9790751;  
 RA Whitmore S.A., Settassian C., Crawford J., Lower K.M., McCallum B.,  
 RA Seshadri R., Corneliisse C.J., Moerland E.W., Cleton-Jansen A.-M.,  
 RA Tipping A.J., Mathew C.G., Savnio M., Savoia A., Verlander P.,  
 RA Auerbach A.D., Van Berkel C., Pronk J.C., Doggett N.A., Callen D.F.;  
 RT "Characterization and screening for mutations of the growth arrest-  
 specific 11 (GAS11) and Cl6orf3 genes at 16q24.3 in breast cancer.";  
 RL Genomics 52:325-331(1998).

CC -!- POLYMORPHISM: THERE SEEMS TO BE TWO ALLELES OF THIS PROTEIN.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF050081; AAC69521.1; -;  
 DR EMBL; AF050080; AAC69520.1; -;  
 DR Genew; HGNC:1197; Cl6orf3.  
 DR MIM; 605179; -;  
 KW Repeat; Polymorphism.  
 FT VARIANT 47 54 MISSING (IN SHORT ISOFORM).  
 FT AS REGA 125 AA; 12619 MW; ACCA40DBE0A79B4 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WTLK 10  
 ||||  
 Db 22 WTLK 25

## RESULT 10

RS3A\_ARCFU STANDARD; PRT; 211 AA.  
 ID RS3A\_ARCFU  
 AC O27964;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 30S ribosomal protein S3ae.  
 GN RPS3AE OR AF2320.  
 OS Archaeoglobus fulgidus.

```

CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
CC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE S3AE FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000944; AAB88936.1; -.
CC TIGR; AF2320; -.
CC InterPro; IPR001593; Ribosomal S3AE.
CC Pfam; PF01015; Ribosomal_S3AE; 1.
CC ProDom; PD003035; Ribosomal_S3AE; 1.
CC PROSITE; PS01191; RIBOSOMAL_S3AE; FALSE_NEG.
CC Ribosomal protein; Complete proteome.
CC SEQUENCE 211 AA; 24334 MW; 3463C7042B0F2796 CRC64;
CC -----
CC
CC Query Match 71.4%; Score 25; DB 1; Length 211;
CC Best Local Similarity 100.0%; Pred. No. 1.4e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 7 WTLK 10
CC ||||
CC Db 14 WTLK 17
CC -----
CC
CC RESULT 11
CC VB4_CRPVK STANDARD; PRT; 212 AA.
CC AC P03124;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
CC DE Probable E4 protein.
CC GN E4.
CC OS Cotton-tail rabbit (shope) papillomavirus (strain Kansas) (CRPV).
CC OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC OC Papillomavirus.
CC OX NCBI_TaxID=31553;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=85166175; PubMed=2984661;
CC RA Giri I., Danos O., Yaniv M.;
CC RT "Genomic structure of the cottontail rabbit (Shope) papillomavirus.";
CC Proc. Natl. Acad. Sci. U.S.A. 82:1580-1584(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02708; -; NOT_ANNOTATED_CDS.
CC PIR; A03674; W4WLRB.
CC Early protein.
CC SQ SEQUENCE 212 AA; 24104 MW; 9FB8DA3FB14ADB5B CRC64;
CC
CC Query Match 71.4%; Score 25; DB 1; Length 212;
CC Best Local Similarity 100.0%; Pred. No. 1.4e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 7 WTLK 10
CC ||||
CC Db 73 WTLK 76
CC -----
CC
CC RESULT 12
CC Y363_RICPR STANDARD; PRT; 218 AA.
CC ID Y363_RICPR
CC AC Q9ZDG6;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein RP363.
CC GN RP363.
CC OS Rickettsia prowazekii.
CC OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
CC OC Rickettsiaceae; Rickettsieae; Rickettsia.
CC OX NCBI_TaxID=782;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Madrid E.
CC RX MEDLINE=99039499; PubMed=9823893;
CC RA Andersson S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
CC RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
CC RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
CC RT "The genome sequence of Rickettsia prowazekii and the origin of
CC RT mitochondria.";
CC RL Nature 396:133-140(1998).
CC CC -!- SIMILARITY: SOME, TO R.PROWAZEKII RP364.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ235271; CAA14822.1; -.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 218 AA; 25780 MW; 2D8BDD7DB938F954 CRC64;
CC
CC Query Match 71.4%; Score 25; DB 1; Length 218;
CC Best Local Similarity 100.0%; Pred. No. 1.4e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 7 WTLK 10
CC ||||
CC Db 46 WTLK 49
CC -----
CC
CC RESULT 13
CC MSRA_BOVIN STANDARD; PRT; 233 AA.
CC ID MSRA_BOVIN
CC AC P54149;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Peptide methionine sulfoxide reductase (EC 1.8.4.6) (Protein-
CC DE methionine-S-oxide reductase) (Peptide Met(O) reductase).

```

```

GN MSRA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal medulla;
RX MEDLINE=96312545; PubMed=8700890;
RA Moskowitz J., Weissbach H., Brot N.;
RT "Cloning the expression of a mammalian gene involved in the reduction
RT of methionine sulfoxide residues in proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2095-2099(1996).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=20519025; PubMed=11063566;
RA Lowther W.T., Brot N., Weissbach H., Matthews B.W.;
RT "Structure and mechanism of peptide methionine sulfoxide reductase, an
RT 'anti-oxidation' enzyme.";
RL Biochemistry 39:13307-13312(2000).
CC -!- FUNCTION: HAS AN IMPORTANT FUNCTION AS A REPAIR ENZYME FOR
CC PROTEINS THAT HAVE BEEN INACTIVATED BY OXIDATION. CATALYZES THE
CC REVERSIBLE OXIDATION-REDUCTION OF METHIONINE SULFOXIDE IN PROTEINS
CC TO METHIONINE.
CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC protein L-methionine S-oxide + reduced thioredoxin.
CC -!- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U37150; AAC48539.1; -.
DR PDB; 1FVA; 08-NOV-00.
DR PDB; 1FVG; 08-NOV-00.
DR InterPro; IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRFAMs; TIGR00401; msra; 1.
DR Oxidoreductase; 3D-structure.
KW Oxidoreductase; 3D-structure.
SQ SEQUENCE 233 AA; 25846 MW; 1F69D612723FBABA CRC64;

Query Match 71.4%; Score 25; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 7 WTLK 10
DB 81 WTLK 84

RESULT 14
PSE2_RAT
ID PSE2_RAT STANDARD; PRT; 238 AA.
AC Q63798;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta
DE subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease
DE subunit 2) (11S regulator complex beta subunit) (REG-beta).
GN PSME2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=95309399; PubMed=7789512;
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,
RA Chung C.H., Shibamura N., Willy P.J., Mott J.D., Slaughter C.A.,
RA DeMartino G.N.;
RT "Primary structures of two homologous subunits of PA28, a gamma-
RT interferon-inducible protein activator of the 20S proteasome.";
RL FEBS Lett. 366:37-42(1995).
CC -!- FUNCTION: Implicated in immunoproteasome assembly and required for
CC efficient antigen processing. The PA28 activator complex enhances
CC the generation of class I binding peptides by altering the
CC cleavage pattern of the proteasome.
CC -!- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC
CC RING.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE PA28 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D45250; BAA08207.1; -.
DR HSP; Q06323; JAVO.
DR InterPro; IPR003185; PA28_alpha.
DR InterPro; IPR003186; PA28_beta.
DR Pfam; PF02251; PA28_alpha; 1.
DR Pfam; PF02252; PA28_beta; 1.
DR Proteasome; Interferon induction.
KW Proteasome; 238 AA; 26857 MW; AFA0F013CECE1CD3 CRC64;

Query Match 71.4%; Score 25; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 7 WTLK 10
DB 111 WTLK 114

RESULT 15
PSE2_HUMAN
ID PSE2_HUMAN STANDARD; PRT; 239 AA.
AC Q9UL46; Q15129;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteasome activator complex subunit 2 (Proteasome activator 28-beta
DE subunit) (PA28beta) (PA28b) (Activator of multicatalytic protease
DE subunit 2) (11S regulator complex beta subunit) (REG-beta).
GN PSME2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99216442; PubMed=10199920;
RA McCusker D., Wilson M., Trowsdale J.;
RT "Organization of the genes encoding the human proteasome activators
RT PA28 alpha and PA28 beta.";
RL Immunogenetics 49:438-445(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95309399; PubMed=7789512;
RA Ahn J.Y., Tanahashi N., Akiyama K., Hisamatsu H., Noda C., Tanaka K.,
RA Chung C.H., Shibamura N., Willy P.J., Mott J.D., Slaughter C.A.,
RA DeMartino G.N.;
RT "Primary structures of two homologous subunits of PA28, a gamma-
RT interferon-inducible protein activator of the 20S proteasome.";
RL FEBS Lett. 366:37-42(1995).

```

```
[3]
RN  SEQUENCE FROM N.A.
RC  TISSUE=Pancreas;
RA  Strausberg R.;
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC  !- FUNCTION: Implicated in immunoproteasome assembly and required for
CC  efficient antigen processing. The PA28 activator complex enhances
CC  the generation of class I binding peptides by altering the
CC  cleavage pattern of the proteasome.
CC  !- SUBUNIT: HETERODIMER OF PSME1 AND PSME2, WHICH FORMS A HEXADIMERIC
CC  RING.
CC  !- INDUCTION: BY INTERFERON GAMMA.
CC  !- SIMILARITY: BELONGS TO THE PA28 FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF079558; AAF02218.1; -.
DR  EMBL; DA5248; BAA08205.1; -.
DR  EMBL; BC004368; AAH04368.1; -.
DR  HSSP; Q06323; LAVO.
DR  PHCI-2DPAGE; Q15129; -.
DR  Genew; HGNC:9569; PSME2.
DR  MIM; 602161; -.
DR  InterPro; IPR003185; PA28_alpha.
DR  InterPro; IPR003186; PA28_beta.
DR  Pfam; PF02251; PA28_alpha; 1.
DR  Pfam; PF02252; PA28_beta; 1.
KW  Proteasome; Interferon induction.
FT  CONFLICT 229 229 N> T (IN REF. 2).
SQ  SEQUENCE 239 AA; 27361 MW; 97A29583AA78A87C CRC64;

Query Match      71.4%; Score 25; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 WTLK 10
    ||||
Db  112 WTLK 115

Search completed: January 29, 2003, 10:45:07
Job time : 6.36364 secs
```